

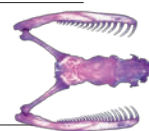
# THIS WEEK

## EDITORIALS

**ASBESTOS** No excuse for countries failing to ban this killer mineral **p.868**

**WORLD VIEW** Clinical pharmacology is dead. Long live TMAP **p.869**

**SNAKE FOOD** Southpaw snails prove hard to swallow **p.870**



## Response required

*Blogs and online comments can provide valuable feedback on newly published research. Scientists need to adjust their mindsets to embrace and respond to these new forums for debate.*

**Y**ou may have seen claims that scientists at NASA have discovered a bacterium that can replace the phosphorus in its DNA with arsenic. You may have heard that this could help the hunt for aliens. You may even have heard that the 'arsenic bacterium' is itself an alien. What you will not have seen or heard is a detailed response from NASA and the scientists involved to online criticism of their work. In the face of worldwide attention on their paper (F. Wolfe-Simon *et al.* *Science* doi:10.1126/science.1197258; 2010), which NASA and the team deliberately courted, the researchers have stuck their heads in the digital sand.

In response to the arsenic bacterium claims, bloggers and researchers raised serious and thoughtful reservations about the paper's methodology and findings. But the authors say that they will not engage with these critics, or with science journalists drawn to the controversy, because such discussion should be moderated in the peer-reviewed literature. Meanwhile, they are urging other scientists to work to replicate their results — a process that will take many months. "We are not going to engage in this sort of discussion," Felisa Wolfe-Simon, the paper's lead author and a NASA astrobiology research fellow at the US Geological Survey in Menlo Park, California, told one *Nature* reporter, "Any discourse will have to be peer-reviewed in the same manner as our paper was, and go through a vetting process so that all discussion is properly moderated."

Purists who hold peer review as the casting vote in such debates will read her words with approval. But the problem is that Wolfe-Simon's reticence is the polar opposite of the fanfare with which NASA trailed her discovery to the public. In an advance press advisory on 29 November, NASA trumpeted an "astrobiology finding that will impact the search for evidence of extraterrestrial life". At a press conference to coincide with the paper's publication, the authors reported a more down-to-Earth, but nonetheless radical, discovery, claiming that an arsenic-tolerant bacterium had rewritten the rules of life as we know them.

Such claims were always likely to bring intensive scrutiny, especially as many scientists think that NASA has form for making extravagant claims in the field of astrobiology. Within two days of the paper appearing, Rosie Redfield, a microbial geneticist at the University of British Columbia in Vancouver, Canada, published a long and detailed critique of what she described as the paper's methodological shortcomings on her blog ([go.nature.com/ddesjw](http://go.nature.com/ddesjw)). She was one of several researchers who used their blogs to question whether the paper's data supported its claims. It was at this point that the authors, previously happy to promote their findings, refused to answer further questions and retreated behind the walls of peer review.

Formal peer review does give criticized authors time to think critically and carefully, and it is a good way to filter out rubbish. But in this case, much of the criticism was already coming from the researchers' peers. And it should be remembered that peer review as conducted by journals

is itself full of differing opinions, and is not the only way to crystallize truth from such disputes. In this instance, a prompt and explicitly provisional response from the authors would have been a better approach, particularly given the way they encouraged the original attention.

*Nature* strongly encourages post-publication discussion on blogs and online commenting facilities as a complement to — but not a substitute for — conventional peer review. Yet it is true that so far

**"Bloggers have an important part to play in the assessment of research findings."**

online commenting and blogs have generally contributed little. Of the thousands of papers published every year, only a few attract substantive comments. And, regrettably, it seems that even those meagre comments rarely spark debate: a study of medical articles in the *BMJ* last August found that few authors bothered to respond to online criticisms of

their papers (P. C. Gøtzsche *et al.* *Br. Med. J.* **341**, c3926; 2010).

Bloggers and online commentators have an important part to play in the assessment of research findings, and many researchers' blogs, in particular, contain better analyses of the true significance of a scientific finding or debate than is seen in much of the mainstream media. Science journalists who repeated NASA's claims on the arsenic bacterium and did not tap into the widespread criticisms, did little to defend themselves from claims of reporting by press release. Blogging scientists, meanwhile, should remember that such informal forums do not excuse insults and casual discourtesy towards colleagues — especially those being urged to respond.

In the end, the scientific truth will prevail, as it usually does. In the meantime, researchers must accept some harsh truths about the speed and spread of digital criticism. ■

## Great expectations

*If Europe's new states are to follow the research roadmap, capacity is as essential as funding.*

**T**o win a national bid to host a new European research facility is, for academics, akin to being chosen to hold the Olympic Games. The warm glow of prestige is matched by the flow of hard cash to regenerate land and communities, while the rush of the best scientific minds to the new equipment can give a major boost to national research performance.

So the Czech Republic, Hungary and Romania are rightly proud to have beaten France and the United Kingdom to jointly host the €800-million (US\$1-billion) Extreme Light Infrastructure (ELI), a

consortium of three independent laser facilities to deliver images at the atomic level.

The project is part of a roadmap for European research infrastructure — a wish list of research facilities drawn up by the best scientific minds across the European Union (EU) — and the first to be built in newer, and often less-well-off, member states.

The ELI is on track to begin construction early next year, but the real test starts now. To build it, the host countries will use EU structural funds — a multi-billion-euro pot established to help narrow the economic and social disparities between member states. Earlier this year, Máire Geoghegan-Quinn, European Commissioner for Research, Innovation and Science, said she hoped to divert €86 billion of EU structural funds to building Europe's "knowledge economy", including research infrastructure. In the past, it has been difficult to track how countries have spent such structural funds, and this lack of transparency has led to a sense of mistrust. As a result, policy wonks in established member states are questioning the merits of using structural funds to support research in Europe, such as on the ELI.

Poland is a major beneficiary of structural funding for research infrastructure, and has been allocated €1 billion over the period 2007–2013. Critics of the approach were handed ammunition earlier this year, when Poland invited a panel of international scientists to assess the research infrastructure it wants to build in the future, partly using structural funds. The country should be applauded for its scientifically responsible approach. But some of the experts on the panel have some concerns about the scientific quality of the country's proposals.

Some projects look more like plans to create networks between national universities, they say, or attempts to build and strengthen national industries, rather than to develop cutting-edge research infrastructures. One project aims to build a knowledge alliance between several universities to help develop foundry and metallurgy industries, but contains no ideas for what research would be conducted in this area. Instead, it focuses on how the institutions can be linked up easily, sited as they are along major highways.

Out of a total of 60 points that each proposal could be awarded, the

highest mark was 45.3; the majority of projects came in at just over half marks. As one scientist on the assessment panel (from a research-intensive member state) commented, only projects awarded the equivalent of 54 points or more would be considered for funding in their home nation. There are also widespread concerns in Europe that the new member states lack the experience to manage large infrastructure projects, including handling budgets, procurement and legal aspects. Insiders at the ELI say that this lack of experience is beginning to show, in preparing accounts for example.

The European Research Advisory Board, an independent advisory committee to the European Commission, echoes these fears in a report published in October. The board is concerned that the power given to member states, to decide which projects to fund with structural funds, directs investment towards building national capacity, rather than cutting-edge research.

**"Some Polish research infrastructure projects look like attempts to build national industries."**

The board recommends that some of the structural funds be held back in a central pot, to be allocated to projects judged to be of a high standard by experts, and which would serve pan-European needs. Although this approach may be better for research as a whole, it doesn't address the difficulties faced in the new member states.

These difficulties are not confined to the newer member states, as those countries involved in building ITER, the fusion test reactor struggling to life near Cadarache, France, have learnt the hard way. Legal and managerial expertise that is crucial to make such projects work must be actively sought and shared. For example, the European Investment Bank's initiative to help new member states prepare financial proposals for major projects could be extended to see projects through to later stages. And a portion of structural funds earmarked for research infrastructure could be set aside to train scientists as managers.

Structural funds for research infrastructure should continue to flow, but more international support is needed to ensure that the structures built around them are sound. ■

# Asbestos scandal

*Irresponsible policies could cause an epidemic of malignant lung disease.*

**V**iewed through an electron microscope, asbestos fibres look like thin glass straws, some no more than a fraction of a micrometre wide. If inhaled, they penetrate the soft alveoli of the lungs and the membranes that line the chest cavity. And there they stay. Over time, damaged cells can cause a malignant disease called mesothelioma, which often kills people, horribly, less than a year after diagnosis.

Before the widespread industrial use of asbestos began in the late nineteenth century, malignant mesothelioma was unheard of, yet it is now responsible for tens of thousands of deaths around the world every year. After the link between asbestos exposure and the disease was convincingly made in 1960, responsible nations eventually took strong measures to remove the mineral from commercial products and to halt mining and export. Less responsible nations did not; this is a scandal that deserves wider attention.

The United States has still not banned asbestos, despite the millions of dollars spent to clear it from homes and from communities near mines. And Canada has been criticized for plans to expand asbestos mining operations, which export the material to India, Indonesia and the Philippines. Although Canada enforces strict guidelines on asbestos use at home to protect its own people, those in countries to which it

sends the mineral have little or no protection. Asbestos exported from Canada and other countries including Russia, Brazil and Kazakhstan is routinely mixed into building materials and consumer products, prized for the same durability that makes it troublesome for living tissue. Owing to the long time between exposure and the onset of disease, 30 years or more, the asbestos trade in North America and elsewhere is creating an epidemic that may take decades to peak and subside.

The minerals industry has long tried to convince regulators that white asbestos — or chrysotile — is safe when handled properly. It argues that only the already controlled forms — blue and brown asbestos, known collectively as amphibole — are of concern.

To support this, industry advocates point to scientific data and studies. Yet although the relevant literature is a mire of conflicting results, this should not be seen as an endorsement of their position. Rather, it reflects a string of industry-sponsored studies designed only to cast doubt on the clear links between chrysotile and lung disease. These are familiar tactics and several countries, including Britain, have seen through them and made the correct decision to ban all forms of asbestos, all of which have been proven to be carcinogenic in humans.

Meanwhile, researchers are finding new causes for concern with other natural fibrous minerals such as erionite (see page 884). Complacency is the problem. Much of the developed world has seen asbestos removed from public spaces, leaving in many minds a false sense of security. The

public should once again be made aware of the risks associated with exposure to mineral fibres, as well as some man-made fibres. And governments must ban the extraction, processing and use of materials that can cause serious disease. ■

**NATURE.COM**  
To comment online,  
click on Editorials at:  
[go.nature.com/xhunq](http://go.nature.com/xhunq)



## Drug development needs a new brand of science

*We need to break with the past to develop new medicines, says*  
**Garret FitzGerald.** *An interdisciplinary NIH centre points the way.*

Last week, the US National Institutes of Health (NIH) voted to launch a National Center for Advancing Translational Sciences, focusing on translational medicine and therapeutics (TMAT), the growing field that aims to speed therapies from the laboratory to the clinic. NIH director Francis Collins called the decision “momentous” — a “disruptive innovation on an institutional scale” — and I think he is right. Only a translational approach can address the fact that the current model of drug discovery and development is unsustainable. Paradoxically, as we have witnessed a successful revolution in drug discovery, a crisis has emerged in drug development. Targets, and the chemistry needed to probe them, can be selected more rationally than ever — yet more and more candidate drugs are proving expensive failures.

One reason is that too many steps are pursued in specialist isolation, in both academia and industry. Too few people can bridge the translational and interdisciplinary divides. This has led to crucial and expensive mistakes in phase II of drug development — when there is often a failure to see an impact on efficacy, a propensity to ignore risks, or a danger of making errors in dose selection for phase III.

The new NIH centre promises to catalyse a much-needed restructuring of the drug-development process. The centre can foster training by absorbing the Clinical and Translational Science Awards (CTSAs) and their educational infrastructure. This will allow scientists to partner in a modular approach to drug development, in which expertise is drawn from distinct sectors and regions as needed to address particular therapeutic challenges. Furthermore, the broad CTSA-supported programmes and infrastructure — from preclinical science to community outreach — could be harvested to support a more efficient approach to drug development, approval and dissemination.

Why has the need for such a radical change emerged? Thirty years ago, the best clinical pharmacology units housed experts from a range of disciplines. Cell biologists worked side by side with colleagues studying model systems and those involved in mechanistic studies of physiology, disease and drug action in humans and pharmacokinetics. Others were trained in chemistry, statistics and toxicology. Blending these heterogeneous talents fostered what we would now call interdisciplinary science, and, in the context of drug development, T1 translational research.

However, as the economics of academic departments shifted, clinical pharmacology fell from favour. Even the term clinical pharmacology has lost its lustre, and now covers only some of what we need. To attract the best and brightest, we need a new brand, backed by funders, academics and industry. Potential students must perceive the field to be hot.

So what shall we call this interdisciplinary, translational endeavour? It is difficult to imagine

anyone rushing to join something called ‘T1 translational research’. ‘TMAT’, on the other hand, captures the fashion for translation, places the discipline in the heart of medicine and indicates the focus on developing novel therapeutics. Adoption of this term by the NIH follows a training programme in TMAT funded by the UK Wellcome Trust. Now we need to realize the potential of this brand and push the idea more widely.

The NIH centre will signal, both to Congress and the biomedical research community, the intimate connection between fundamental science and the accelerated delivery of cures to the general public. This is not a zero-sum game: success of translation requires investment in basic science. By developing sustainable career structures in TMAT, the centre can reverse the flow of bright young scientists into specialist silos.

Joint investments in training, infrastructure and programmes would ensure that the efforts of the new centre would improve, not compete with, the translational efforts of disease-focused institutes and centres within the NIH.

The new TMAT centre could also act as a visible point of contact for extramural partners, including industry, charitable foundations and the US Food and Drug Administration, to buy into the restructuring required to move to a more modular approach to drug discovery and development. A looser, more distributed model spanning pharma, biotech and academia could then draw on knowledge more easily, and apply it more efficiently.

It is a big challenge, and two particular obstacles come to mind. First, we must revise how we reward ideas. At present, defence of intellectual

property relies on patents on the composition of matter, usually molecules, most of which never become approved drugs. To make sure that they do, many people with diverse skill sets have to work effectively together. Inside a company, it is easy to reward everybody involved. As companies fragment, we should consider new models of intellectual property. Perhaps the financial rewards of a patent should be postponed until a drug is a profitable success — and a formal mechanism found to distribute rewards among all those who helped to make it happen.

Second, we will need common standards of data protection and privacy, and shared infrastructure that allows secure and compliant sharing of diverse types of information, including clinical data, across countries and sectors. This is the foundation upon which a global TMAT enterprise can be established. In some ways, this is the greatest challenge of all, but it can be done. As T. S. Eliot said: “Only those who will risk going too far can possibly find out how far one can go.” ■ **SEE NEWS P.877**

**Garret FitzGerald** is director of the Institute of Translational Medicine and Therapeutics at the University of Pennsylvania in Philadelphia.  
e-mail: [garret@exchange.upenn.edu](mailto:garret@exchange.upenn.edu)

**WE MUST  
REVISE HOW WE  
REWARD IDEAS  
AND WILL NEED  
COMMON  
STANDARDS  
OF DATA PROTECTION.**

➔ **NATURE.COM**  
Discuss this article  
online at:  
[go.nature.com/u9jfyq](http://go.nature.com/u9jfyq)



# RESEARCH HIGHLIGHTS

Selections from the  
scientific literature

## BOTANY

### Final frontier of flowering plants

Half of the world's yet-to-be-discovered flowering plant species may already have been collected, and now languish in herbarium cabinets.

While reclassifying varieties of *Strobilanthes*, a genus of purple-flowered plants from Asia, Robert Scotland of the University of Oxford, UK, and his colleagues noticed that many of the 60 species they described had been collected many years before. This lag ranged from 1 to 210 years and averaged more than 30 years for more than 3,000 species in 6 plant genera, including *Strobilanthes*. Just 16% of these plants were classified within 5 years of discovery.

If this trend holds for other flowering plants, 47% to 66% of the planet's estimated 70,000 undiscovered species are waiting to be unveiled in herbaria.

*Proc. Natl. Acad. Sci. USA*  
doi:10.1073/pnas.1011841108  
(2010)

For a blog entry on this research, see [go.nature.com/vcqjg](http://go.nature.com/vcqjg).

## ANIMAL BEHAVIOUR

### Caterpillars whistle for safety

When under attack, walnut sphinx caterpillars (*Amorpha juglandis*; pictured),



whistle. An 1868 *Canadian Entomologist* paper, "Musical larvae," first reported these shrieks, but their purpose wasn't clear.

Jayne Yack at Carleton University in Ottawa, Canada, and her team now show that the whistle, produced through openings along the body called spiracles, is a defence against predators. Simulated attacks with blunt tweezers caused the caterpillars to pull their heads back, forcing air through two of the spiracles in a succession of squeaks.

When confronted by

their real predators, yellow warblers, the caterpillars whistled each time the birds swooped in for attack, repelling multiple assaults until the warblers gave up. *J. Exp. Biol.* 214, 30–37 (2011)  
For videos, see [go.nature.com/zgeqyc](http://go.nature.com/zgeqyc).

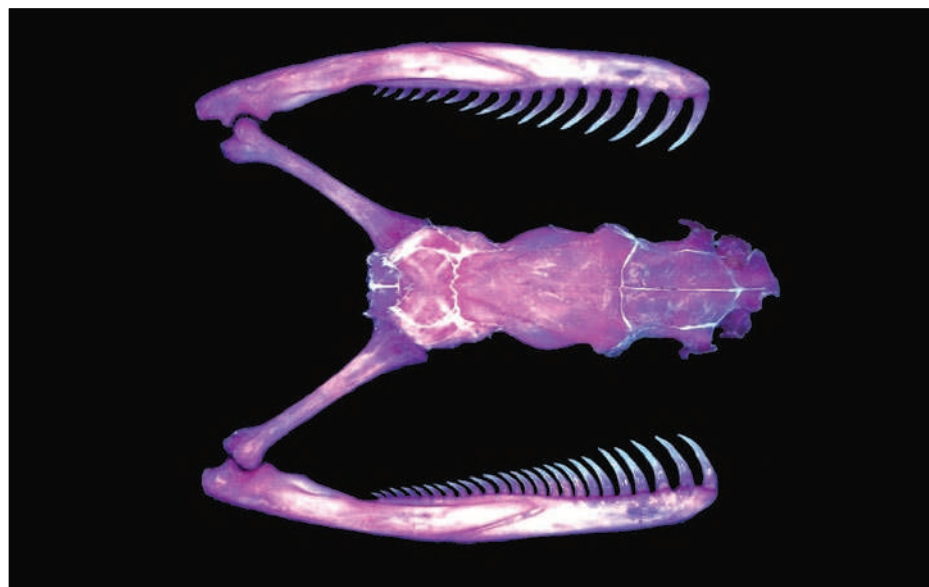
## NANOTECHNOLOGY

### Pressed to breaking point

Every day in labs around the world, a technique using high-frequency sound

waves — ultrasonication — is used to break up carbon nanotubes. But no one really understands the underlying mechanism. Kyung-Suk Kim at Brown University in Providence, Rhode Island, and his collaborators have shed some light on the interplay between nanotubes and the minute bubbles created by the sound waves under water.

When the bubbles implode, tubes in the water near them are suddenly compressed along their lengths. The tubes buckle, and some atoms



## EVOLUTION AND ECOLOGY

### Twisted tale of snail evolution

Dextral snail shells coil rightwards, and sinistral shells coil leftwards. Sinistral *Satsuma* snails cannot mate with right-coiling *Satsuma* species, leading scientists to wonder how sinistrality could have spread through dextral populations.

Masaki Hosoi of Tohoku University in Sendai, Japan, and his colleagues show that sinistrality has arisen independently multiple times in *Satsuma*, and more often where snakes in the Pareatidae family occur.

The team found that the *Pareas iwaskii* snake, which preys on the molluscs, must stick to right-coiling species as its jaws are specialized for grasping them. (Snake jaw, with extra teeth on the lower mandible, pictured.) That gave sinistral individuals an adaptive advantage, allowing left-coiling species to emerge.

*Nature Commun.* doi:10.1038/ncomms1133 (2010)  
For a longer story on this research, see [go.nature.com/9fetev](http://go.nature.com/9fetev).



are knocked off, weakening the tube until it ultimately breaks.

*Proc. R. Soc. A* doi:10.1098/rspa.2010.0495 (2010)

## GENETICS

## Sex and the social slime mold

A single gene is sometimes all it takes to change a slime mold's sexual identity.

The social amoeba *Dictyostelium discoideum* has three different sexes — members of one sex, or 'mating type', can fuse with either of the other two to form giant, dormant cysts. But little is known about what genes determine the sexual identity of a slime mold.

Gareth Bloomfield of the Medical Research Council molecular biology lab in Cambridge, UK, and his colleagues found a region of the *D. discoideum* genome that differed among sexes. Deleting a gene from this region prevented mating-type I from coupling with mating-type II; reintroducing the gene restored normal sexual orientation. Meanwhile, swapping sex genes from one mating type to another caused the amoebae to switch sexual partners.

*Science* 330, 1533–1536 (2010)

## DEVELOPMENT

## Mother's dinner, daughter's nose

The smell of mouse mothers' food influences the olfactory anatomy of their pups, and primes them to prefer the same flavours as their mothers.

Josephine Todrank at the University of Colorado, Denver, and her colleagues studied lines of mice in which select olfactory sensory neurons that responded to smells such as cherry or mint were tagged with the gene for green fluorescent protein. The mothers were given scented food while either gestating or nursing their

litters, or during both phases. When their pups were tested at 20 days old, fluorescence revealed larger glomeruli — bundles of synapses — formed by neurons specific to the smells added to their mother's food. Pups also preferred the smells of the food their mothers ate.

Such preferences could predispose animals to choose familiar and safe foods, although in humans they could backfire to plant the seed of preference for alcohol or unhealthy foods, the authors say.

*Proc. R. Soc. B* doi:10.1098/rspb.2010.2314 (2010)

## ORGANIC ELECTRONICS

## Currency circuitry

Modern anti-counterfeiting features on banknotes are getting more sophisticated, ranging from complex and colourful watermarks to holograms and foil strips. Now Ute Zschieschang of the Max Planck Institute for Solid State Research in Stuttgart, Germany, and her colleagues have added yet another weapon to the



arsenal: trackable digital circuits.

The researchers fabricated low-voltage organic transistors on the surface of a €5 note (pictured), using a 3-nanometre-thick insulating layer made of aluminium oxide and octadecylphosphonic acid that could be deposited without damaging the surface of the banknote. A total of 92% of the deposited transistors were functional — a high enough proportion for

## COMMUNITY CHOICE

The most viewed papers in science

## ORGANIC CHEMISTRY

## Making maoecrystal V

HIGHLY READ  
in November on  
pubs.acs.org

The complex molecule maoecrystal V has been synthesized in the laboratory, after six years of intense effort by high-profile chemists. Zhen Yang and his colleagues at Peking University in Beijing created the sought-after compound — which shows potent activity against cancer cells — in a concise 16-step synthesis. It was originally extracted from a Chinese herb (*Isodon eriocalyx*) that has long been used as a folk medicine to treat flu and inflammation, and has already produced a number of potential anticancer agents. By varying the laboratory synthesis, chemists will be able to make and test closely related structures that may prove better medicines than maoecrystal V itself.

*J. Am. Chem. Soc.* 132, 16745–16746 (2010)

the circuits to work reliably.

*Adv. Mat.* doi:10.1002/adma.201003374 (2010)

## GENETIC TESTING

## It's never too early to get sequenced

A developing baby's entire genome is hidden in its mother's blood, potentially offering a non-invasive test for congenital diseases. Dennis Lo of the Chinese University of Hong Kong and his colleagues sequenced billions of DNA base pairs from the plasma of a pregnant woman and then developed a way to distinguish her DNA sequences from the fetus's.

Both parents carried a single mutation for  $\beta$ -thalassaemia, a rare blood disorder caused by two faulty copies of the gene *HBB*. Lo's analysis demonstrated that the father had passed on his mutation, but the mother had given the fetus a healthy copy of *HBB*, sparing it from  $\beta$ -thalassaemia. Such genetic screening could replace invasive prenatal diagnostic tests such as amniocentesis.

*Sci. Transl. Med.* 2, 61ra91 (2010)  
For a longer story on this research, see go.nature.com/djxvga.

## PLANETARY SCIENCE

## Impacts sent bling to early Earth

Call it a gift. Late in the Solar System's formation, a shower of objects up to the size of Pluto delivered to Earth a large quantity of rock containing gold, platinum and other elements that bind readily with iron. Researchers believe these elements were added to the mantle late in Earth's development, because if they had been present when the planet was molten, they would have sunk to its core, with iron.

William Bottke of the Southwest Research Institute in Boulder, Colorado and his colleagues used abundances of iron-loving elements on Earth, the Moon and Mars to model how later impacts from large objects could have replenished reserves in the planets' mantles. The findings may also explain the sizes of the oldest craters on the Moon and Mars.

*Science* 330, 1527–1530 (2010)  
For a longer story on this research, see go.nature.com/bbdewm.

## NATURE.COM

For the latest research published by Nature visit:

[www.nature.com/latestresearch](http://www.nature.com/latestresearch)

# SEVEN DAYS

The news in brief

## RESEARCH

### Crop catalogue

A global search to gather the wild relatives of essential food crops such as wheat, barley and rice has been launched by the Global Crop Diversity Trust, based in Rome. The ten-year initiative, announced on 10 December, aims to increase food security by finding genetic traits that might be suited to future climates. Samples of wild plants will now be conserved alongside existing stores of domesticated seeds (such as the Svalbard Global Seed Vault on the Norwegian island of Spitsbergen). See [go.nature.com/l8mgn2](http://go.nature.com/l8mgn2) for more.

### Higgs hunt extended

A 15-month shutdown to upgrade the Large Hadron Collider is set to be delayed by a year to the end of 2012. The extended run will be used by scientists at the particle-physics laboratory CERN near Geneva, Switzerland, to hunt for the elusive Higgs particle at the collider's current collision energies. The plan is likely to be agreed by CERN's management and council in January. See page 876 for more.

### African innovation

Africa is struggling to turn local discoveries into drugs and other health-care inventions, according to papers produced by the McLaughlin-Rotman Center for Global Health in Toronto, Canada. The reports, published by BioMed Central on 13 December, identify 25 'stagnant technologies' languishing in African health-care institutions, including several drug candidates and a dipstick test for schistosomiasis. Scientists have no incentive to commercialize results, there is scant institutional support for knowledge transfer, and



C. THOMPSON/SPACEX

## Private spaceflight success

SpaceX (Space Exploration Technologies Corporation) has become the first private firm to launch a spacecraft into orbit and return it to Earth. On 8 December, its reusable 'Dragon' capsule was launched on a Falcon 9 rocket from Cape Canaveral, Florida. Completing two orbits, it splashed down in the Pacific Ocean

roughly 800 kilometres west of Mexico. NASA expects the craft to ferry astronauts, supplies and research materials to the International Space Station when its shuttle fleet retires next year. SpaceX, based in Hawthorne, California, hopes to dock Dragon with the station during its next demonstration launch, scheduled for 2011.

existing regulatory frameworks inhibit innovation, the papers say. See [go.nature.com/py46rh](http://go.nature.com/py46rh) for more.

### Venus probe flop

In a bitter disappointment for Japan's space agency, its Akatsuki spacecraft failed to enter orbit around Venus on 6 December. The probe was intended to monitor the hot planet's atmosphere, but must now wait six years for another chance to reach orbit. See page 882 for more.

## POLICY

### NIH access

A key panel of advisers to the US National Institutes of Health (NIH) voted last week to open the Clinical Center — the agency's huge research hospital in Bethesda, Maryland

— to outside investigators. The Scientific Management Review Board on 7 December recommended extramural scientists be given access to the facility, where roughly 1,500 patient studies are in progress at any given time (see *Nature* **466**, 172; 2010). The same board voted to establish a translational-medicine centre at the NIH (see page 877 for more).

### European patent

Countries in the European Union (EU) have broken through a decade-long impasse over establishing a low-cost single European patent system. At a meeting of the EU competitiveness council on 13 December, 11 countries agreed on a plan to translate EU patents into English and one of French or German;

12 others suggested that they would join the proposal. Italy and Spain voted against the scheme, but countries invoked an 'enhanced cooperation' provision, which allows them to progress without attaining unanimous agreement. A common European patent could be in place by the end of next year; a formal decision is expected in March.

### Anthrax report

The US National Academy of Sciences has delayed releasing a long-awaited report on the investigation into the 2001 anthrax attacks, after a request by the Federal Bureau of Investigation (FBI). The report examines the scientific evidence used by the FBI to accuse microbiologist Bruce Ivins of the attacks, which killed five people.

S. ZAKLIN/GETTY IMAGES

Ivins committed suicide in 2008. After seeing a draft copy, the FBI said hundreds more pages of previously undisclosed documents should be considered by the investigation, which will now continue until February 2011.

## Cancún climate deal

United Nations climate talks in Cancún, Mexico, ended with an agreement by developed and developing countries to reduce greenhouse-gas emissions — largely approving commitments made in last year's Copenhagen Accord. See page 875 for more.

### PEOPLE



## Committee chairs

Ralph Hall (Republican, Texas) was on 8 December confirmed as the new chairman of the US House Committee on Science and Technology. Hall (pictured) has made it clear that he will take a hard line against attempts to regulate greenhouse gases. Fred Upton (Republican, Michigan)

— who has frequently supported environmental legislation — will chair the House Committee on Energy and Commerce.

## NASA chief scientist

Waleed Abdalati will be NASA's chief scientist from 3 January, the agency's administrator Charles Bolden announced on 13 December. A researcher on polar ice who worked at NASA for a decade until 2008, Abdalati is currently director of the Earth Science and Observation Center at the University of Colorado at Boulder. He is NASA's first chief scientist since James Garvin, who served in the post during 2004–05.

## Nobel chemist dies

John Fenn (pictured), who shared the 2002 Nobel Prize in Chemistry, died on 10 December aged 93. In the late 1980s, he developed electrospray ionization, a way to gently separate clumped proteins into a fine spray of individual molecules. This method, when combined with mass spectrometry, gave scientists a tool to quickly identify proteins via their mass and helped to launch the field of proteomics. In 2005, Fenn lost a legal battle over the patent rights to Yale University in New Haven, Connecticut, where he developed the technique.



He had moved to Virginia Commonwealth University in Richmond in 1994.

### BUSINESS

## Biotech bid

On 8 December, US pharmaceutical giant Johnson & Johnson issued a long-awaited public offer to buy Crucell, a biotechnology firm headquartered in Leiden, the Netherlands. Johnson & Johnson, in New Brunswick, New Jersey, offered to pay €1.75 billion (US\$2.3 billion) for Crucell, which specializes in vaccines and antibody therapies. Crucell's board of directors unanimously supports the deal, and shareholders will vote on the matter on 8 February.

## Cheap sequencing

Research-services giant Life Technologies of Carlsbad, California, announced on 14 December that it is now selling benchtop DNA sequencers to labs for less than US\$50,000. The

sequencers, the first of a new wave of affordable machines to reach the market, were developed by start-up firm Ion Torrent in Guilford, Connecticut, which was bought by Life Technologies in August. Current sequencing technologies label nucleotides with dyes, but the new machine uses semiconducting chips to detect hydrogen ions released as nucleotides are added to a DNA strand. Life Technologies also announced three out of seven \$1-million prizes that it will award for solving key challenges in low-cost DNA sequencing. See [go.nature.com/mbhs6a](http://go.nature.com/mbhs6a) for more.

## TB diagnosis

The World Health Organization (WHO) said on 8 December that a test that can rapidly diagnose tuberculosis (TB) was a 'major milestone' for disease control. The DNA-based 'Xpert MTB/RIF' test, developed by the non-profit Foundation for Innovative New Diagnostics in Geneva, Switzerland, and the company Cepheid, based in Sunnyvale, California, can detect TB in around 100 minutes. Traditional tests, based on sputum-smear microscopy, can take up to three months to yield results, the WHO said. The new test is costlier, so will need donor funding, although Cepheid will cut prices by 75% for poorer nations.

S. J. FERRELL/CONGRESSIONAL QUARTERLY/GETTY IMAGES

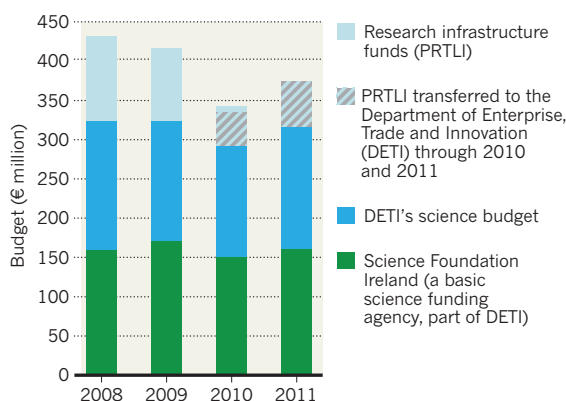
SOURCE: DETI/SFI

## TREND WATCH

After last year's drop in research-grant funding, the Irish government kept a promise to spare researchers further pain in its austerity budget for 2011–14, announced on 7 December. Total funding for basic science has flatlined, not including inflation, since 2008. But the Department of Enterprise, Trade and Innovation announced a 12.5% increase in its science and technology budget compared with 2010. The basic-science funding body Science Foundation Ireland saw a 7% increase in its share.

### IRELAND'S SCIENCE BUDGET

Despite cutting €6 billion (US\$8.1 billion) from its budget, Ireland maintained its science funding.



### CORRECTION

The story 'Synchrotron cuts' (Nature **468**, 736; 2010) incorrectly gave the three-year budget of the European Synchrotron Radiation Facility as €86.8 million. That number is the facility's annual budget. The brief also said that cuts to "operating time" would be made. To clarify, two existing beam lines will be closed, but the accelerator will continue to run on its normal schedule.

**NATURE.COM**

For daily news updates see:

[www.nature.com/news](http://www.nature.com/news)



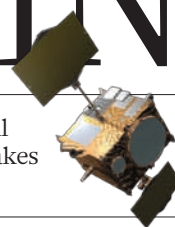
# NEWS IN FOCUS

**PHYSICS** Doomwatch denied as LHC fails to generate mini black holes **p.876**

**GENOMES** Recreational ancestry analysis takes to the web **p.880**

**SPACE** Japanese probe misses its date with Venus **p.881**

**CANCER** Hazardous dust raises concerns in Turkey and North Dakota **p.884**



Christiana Figueres (left), executive secretary of the UN climate framework, speaks with Mexico's foreign minister Patricia Espinosa during the recent climate-change conference in Cancún.

## ENVIRONMENTAL SCIENCE

# Last-minute deal saves climate talks

*Modest agreement brings developing world into the field.*

BY JEFF TOLLEFSON

International negotiators did what they needed to do in Cancún, Mexico, to keep the United Nations climate talks from collapsing into the failure that many had feared. The true extent of their success, however, will depend on what comes next.

Working into the small hours of 11 December, negotiators agreed that both developed and developing countries will act to reduce greenhouse-gas emissions — and that those actions

will be registered and subjected to some form of international verification. The accord represents a major shift for developing countries, which faced no such commitments under the existing Kyoto Protocol, due to expire in 2012. The conference also reached a historic agreement on forest protection, and advanced programmes to help the developing world adopt clean energy and adapt to climate change.

“It’s not just about process, it’s about substance,” said Connie Hedegaard, the European Union’s top climate official, in a news conference

as the all-night talks wrapped up. “We have proven that multilateralism can create results.”

The negotiators built on the broader framework of last year’s Copenhagen Accord. “Ideas that were just skeletal last year are now approved and elaborated,” says Todd Stern, the chief US climate negotiator. But unlike the Copenhagen document, which was blocked by a few countries in a rowdy final session, the Cancún agreement was adopted unanimously. “It’s a significant step forward,” says Stern.

The agreement sets a goal of limiting average warming to 2°C above preindustrial levels, while acknowledging that current commitments registered under the UN climate framework do not add up to meeting that goal. It also pledges to periodically review the goal on the basis of “the best available scientific knowledge”. Parties agreed to help mitigate emissions according to their own capabilities while a new international reporting system tracks their progress; the burden on developing countries would mainly fall on rapidly emerging economies.

David Victor, director of the Laboratory on International Law and Regulation at the University of California, San Diego, calls the global nature of the agreement “tremendously important”, although it mostly amounts to registering previous progress. “They have managed to reach agreement by moving the goalposts closer to the ball,” he says.

The talks nearly faltered over the future of the Kyoto Protocol, which requires the world’s wealthiest nations — apart from the United States, which has never signed on — to meet specific emissions targets. Japan made waves at the meeting’s outset by announcing it would not support an extension of the treaty after it expires. The final text of the Cancún agreement defers the Kyoto question for another year.

The other major dispute in Cancún was between the world’s two largest greenhouse-gas emitters — the United States and China. With a little help from India, which staked out a position in the middle and softened its rhetoric against binding commitments, both parties were able to agree on some basic requirements for reporting and verifying climate pledges.

Progress on these issues was crucial to keep the talks alive and to open the way to more focused decisions in other areas. First among them was the deforestation agreement, which ▶

➔ **NATURE.COM**  
For our full coverage of the Cancún meeting, visit:  
[go.nature.com/9ualzj](http://go.nature.com/9ualzj)

► advocates hailed as a major accomplishment that will bolster bilateral and multi-lateral efforts already under way.

The Cancún agreement establishes a framework that would allow wealthy nations to pay others for “reducing emissions from deforestation and forest degradation” (known as REDD) and augmenting the carbon stocks locked up in forests. Collectively, the programme is called REDD-plus. The agreement requires developing countries to craft a national plan, establish a baseline for historic emissions from forest loss and create a system for monitoring their forests. Just as importantly, says John Niles, director of the Tropical Forest Group in San Diego, California, the agreement calls on an existing technical body to look into the programme rules and requirements and then report back within a year.

“Once we have those requirements, then everybody knows what we have to get to before any money changes hands,” says Niles. “This is the biggest decision we could have asked for.”

**“They have managed to reach agreement by moving the goalposts closer to the ball.”**

Delegates also agreed to establish a Green Climate Fund to be managed by representatives of the developed and developing world to help channel aid; a Cancún Adaptation Framework to help to guide decisions on funding for adaptation measures in the developing world; and a technology-transfer mechanism to supply developing nations with technology for clean energy and adaptation. As promised in Copenhagen, industrialized nations will provide some US\$30 billion for these programmes by 2012, and up to \$100 billion annually by the end of the decade, although where the money would come from remains unclear.

Tim Gore, climate-change adviser for Oxfam International, based in Oxford, UK, lauds the Green Climate Fund but says that countries missed an opportunity to spell out long-term climate funding, perhaps through a levy on international aviation and shipping. Nonetheless, the agreement represents “a solid step”, says Gore. “They are now walking in the right direction, but they need to start running.” ■

## PARTICLE PHYSICS

# No black holes, but extra time at LHC

*Upgrade likely to be delayed in bid to capture Higgs particle.*

BY GEOFF BRUMFIEL

**T**he end of the world is not nigh after all. Flouting predictions from some theorists, microscopic black holes have so far failed to appear inside the Large Hadron Collider (LHC), scientists there have revealed.

The result, which will be posted this week on arXiv.org, comes as researchers make plans to keep the LHC running until the end of 2012, rather than 2011 as previously scheduled. The 27-kilometre collider at the particle-physics laboratory CERN near Geneva, Switzerland, had endured delays and a crippling breakdown before finally surging to life late in 2009, and physicists say it is now performing above expectations.

Predictions of mini black holes forming at collision energies of a few teraelectronvolts (TeV) were based on theories that consider the gravitational effects of extra dimensions of space. Although the holes were expected to evaporate quickly, some suggested that they might linger long enough to consume the planet. But scientists at the Compact Muon Solenoid (CMS) detector now say they found no signs of mini black holes at energies of 3.5–4.5 TeV. Physicist Guido Tonelli, the detector's spokesperson, says that by the end of the next run, the LHC should be able to exclude the creation of black holes almost entirely.

The find is one of a stream of recent papers from the LHC, made possible by the machine's unexpectedly high performance. “We were very surprised by how well behaved the

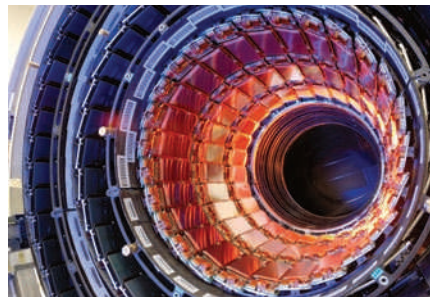
machine was when we started really pushing it to its limit,” says Steve Myers, the CERN physicist who oversaw this year's LHC operations. As a consequence, physicists are increasingly optimistic that they may be able to detect the elusive Higgs boson earlier than expected. The particle, the LHC's best-known quarry, and its associated field are thought to endow other particles with mass.

Initially, physicists were not sure that the LHC could create and detect the Higgs at the machine's current energies, and CERN managers had planned a 15-month hiatus from the start of 2012 for an upgrade that would allow

it to run at higher energies. But a growing consensus holds that even without the upgrade, the LHC will be able to explore most of the energy range in which a standard Higgs particle might be found. Sergio Bertolucci, CERN's director for research and computing, adds that there are political reasons

to extend the run. The world's second most powerful accelerator, the Tevatron at Fermilab in Batavia, Illinois, is nipping at the LHC's heels as it gathers a growing body of data in its own Higgs hunt. Moreover, the potential success of the LHC is likely to influence European plans for high-energy physics, as well as a global plan for a next-generation linear collider. Both face big budget decisions in the next few years.

The plan to extend the LHC's run will be discussed at a meeting of LHC managers in Chamonix, France, in late January, with a final decision expected shortly after. ■



No black holes here: the Compact Muon Solenoid.

M. BRICE/CERN

**MORE ONLINE**

## VIDEO STORY



Ancient Greek computer rebuilt — out of lego  
[go.nature.com/7po7si](http://go.nature.com/7po7si)

## DAILY NEWS

- Why African discoveries languish in labs [go.nature.com/py46rh](http://go.nature.com/py46rh)
- Gut tissue cultured [go.nature.com/9enzfc](http://go.nature.com/9enzfc)
- No evidence of time before Big Bang [go.nature.com/lbwioo](http://go.nature.com/lbwioo)
- Letters defend Nobel laureate against Nazi charges [go.nature.com/oyiyici](http://go.nature.com/oyiyici)

## VIDEO STORY



Sleep-deprived bees make poor waggle dancers  
[go.nature.com/pulztn](http://go.nature.com/pulztn)

J. EVESON/FLPA





C. SONO DEVILLA/GETTY

## Q&A Francis Collins

# The bridge between lab and clinic

*Francis Collins, director of the US National Institutes of Health (NIH), has made the translation of basic research to the clinic a top priority. On 7 December, that goal moved a step closer to realization when the Scientific Management Review Board (SMRB) — an NIH advisory body — voted nearly unanimously to recommend establishing a centre for translational medicine. Collins had told the review board that despite a “dizzying rate” of basic science discoveries, “far too often promising diagnostic devices and treatments are not making it to market”. He urged that the NIH step into the breach. The proposed National Center for Advancing Translational Sciences would combine several existing programmes, have a budget of at least US\$650 million and could be up and running by October 2011.*

### What is the significance of this vote?

This is a momentous occasion. The SMRB has recommended the formation of a new centre based on scientific arguments about opportunities in translation. I think this is a signal moment, placing the NIH in a new position to play a more muscular part in therapeutic development. That being said, it in no way discounts all of the other things that we need to be doing in basic science, nor does it say that all of [NIH-funded] clinical research is going to be folded into this enterprise.

**➔ NATURE.COM**  
Special collection  
on translational  
research:  
[go.nature.com/px8c4m](http://go.nature.com/px8c4m)

### What do you say to concerns that the new centre will draw resources away from basic research?

I want to be very reassuring: although this centre is a new structure, it will not have much of an impact on the overall distribution of funds between basic and clinical research.

### How should basic researchers regard the centre?

I think some basic scientists will be quite excited about the opportunity to be more connected with the clinical benefits of their own discoveries. Not that that's at all necessary or required, or that basic scientists who don't feel

that inclination should be considered somehow unmotivated. Science for science's sake is also a wonderful way to learn about life. But I do think there is more here that is positive than is negative for a basic scientist, if people will step back from their anxieties about budgetary considerations. That being said, we should all be anxious about the overall budget right now with the expectation that dollars for biomedical research are going to be very hard to come by in the next year or two. But that can't be a reason to stop promoting innovation.

### The Clinical and Translational Science Awards, which account for nearly 40% of the budget of the National Center for Research Resources (NCRR), are being moved to the new centre. What happens to the rest of the NCRR's programmes?

Again, the strong assurance is that these programmes are valued, that they will be supported, that the people involved in them are doing great work. There is no intention here to dismantle them. But if there are opportunities to reorganize and reassign these programmes in ways that make them more interactive with what we are trying to do in this new centre, well, that seems like a good thing to consider.

### What do you say to critics who contend that you're just rushing this through?

Well, I'm a guy in a hurry. I will admit it. When I contemplate the urgency of finding treatments and cures for disease, it is hard to be comfortable with an argument that says: 'go slow and take your time'. The SMRB took a comprehensive look at the situation and concluded that the scientific opportunities are here now. Why would you want to delay?

### If this were easy, drug companies would not be struggling with languishing new-drug pipelines. What can the NIH do with this centre that the pharmaceutical giants aren't already doing?

It most certainly will not be easy. But there has been a recent deluge of discoveries about the molecular pathogenesis of disease. This has revealed hundreds of new potential drug targets. For rare and neglected diseases, economic considerations will limit private-sector interest; but NIH-funded researchers can explore the earlier stages in the drug-development pipeline to 'de-risk' projects that would otherwise lie untouched. Similarly, for common diseases, many of the new molecular discoveries are of uncertain value for drug development, but NIH investigators can validate these drug targets and develop promising lead compounds, as well as carrying out process engineering on the pipeline itself. The goal will be to bring each project just far enough to become of interest to the private sector to pick up. ■ [SEE WORLD VIEW P.869](#)

INTERVIEW BY MEREDITH WADMAN



## PHARMACEUTICALS

# Slim spoils for obesity drugs

*Drug makers struggle to find viable treatments for global epidemic.*

BY HEIDI LEDFORD

When an obesity drug that he had helped to invent came up before a panel of US Food and Drug Administration (FDA) advisers last week, physiologist Michael Cowley couldn't bear to watch. "It's like watching your favourite team," says Cowley, director of the Monash Obesity and Diabetes Institute in Victoria, Australia. "You worry that if you pay too much attention they'll lose."

Many thought that Cowley's drug, Contrave, didn't stand a chance. The same panel had already voted against two obesity drugs this year and said a third should be pulled off the market.

But, defying expectations, the panel voted in favour of Contrave, making it the first obesity drug to win a recommendation for approval in more than a decade. It was a rare taste of success for a field that has progressed so slowly that many have abandoned it altogether. "There is surprisingly little activity now given the potential size of the market and the high unmet need," says Michael Hay, an analyst at the market-research firm Sagient Research Systems, based in San Diego, California.

Effective obesity drugs have proven enormously difficult to develop. The brain circuits responsible for appetite overlap with those that control other important functions, including mood, raising the risk of side effects. And obese patients would probably have to take a drug for years, with testing involving large patient populations, driving up development costs.

Most devastating to the field was the failure of drugs designed to block receptors in the brain that respond to appetite-stimulating chemicals

called cannabinoids. Several major pharmaceutical firms had pursued this angle, but gave up after the FDA's advisory panel voted down Paris-based Sanofi-aventis's drug rimonabant in 2007. The London-based European Medicines Agency had already approved rimonabant, but in 2008 it advised doctors not to prescribe it given the risk of suicidal tendencies.

Hopes were running high for a turnaround, with three obesity drugs coming before the FDA this year. First up was a drug called Qnexa from Vivus, based in Mountain View,

in San Diego. But on 16 September, the panel again voted no. The drug simply didn't work well enough to make it worth the safety risks, says Abraham Thomas, an endocrinologist at the Henry Ford Hospital in Detroit, Michigan, who chaired the committee.

By the time Contrave took the stage, market investors had become sceptical (see graph). Contrave, developed by Orexigen Therapeutics, based in La Jolla, California, and co-founded by Cowley, is also a blend of two approved drugs. One, bupropion, is an antidepressant that blocks the effects of the neurotransmitter noradrenaline. The other, naltrexone, inhibits the effects of opioids on the brain and is used to treat alcoholism. Together, the two boost the activity of a brain circuit called the POMC pathway, which reduces hunger. A final decision on Contrave is expected early next year.

But even if approved, Contrave will hardly spell the end of the obesity epidemic. In one recent study, those who took the drug lost only about 8% of their body weight over six months — little better than orlistat, the only over-the-counter obesity drug currently available in the United States.

Most US health-insurance plans don't cover weight-loss drugs, making a minimally effective drug even less appealing.

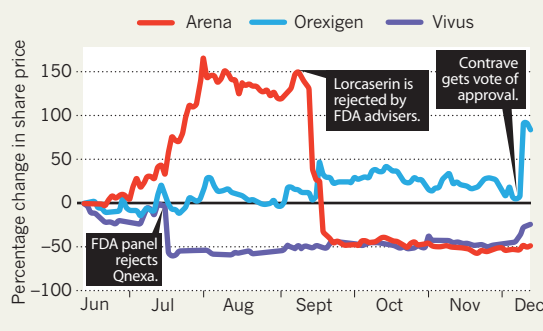
A change of fortune may require a change of strategy, says Thomas Hughes, chief executive of Zafgen, an obesity drug company in Cambridge, Massachusetts. In the past, most companies pursued drugs that would restrict appetite. Now, he says, they're looking for new ideas, "but those ideas are few and far between".

The next obesity drugs likely to face the FDA target metabolism, says Hay (see table). Victoza (liraglutide), developed by Novo Nordisk, based in Bagsvaerd, Denmark, to treat type 2 diabetes, mimics a gut hormone called 'glucagon-like peptide 1', which boosts insulin sensitivity and slows stomach emptying.

Although more than a dozen other drugs are in early clinical trials, Thomas remains pessimistic. Many of them target pathways that have already been tried, he notes, or are reformulations of drugs that didn't make it the first time around. "You can see what's in phase I clinical trials now and the answer is 'nothing very exciting'," agrees Stephen Bloom, an endocrinologist at Imperial College London. "The obesity problem is unsolved, and looks like it's going to stay that way for quite some time." ■

## MARKET MOVEMENTS

Only one of a trio of obesity drugs submitted to a Food and Drug Administration panel got the green light.



California. Qnexa is a combination of two drugs already on the market. One, phentermine, is a brain stimulant that suppresses appetite. The other, topiramate, is a treatment for epilepsy. On 15 July, FDA advisers said that the drug's modest weight-loss benefit was not sufficient to counterbalance the risk of its side effects, which include memory problems and birth defects.

The next drug to go before the panel was lorcaserin, made by Arena Pharmaceuticals

## UP-AND-COMING: THE NEXT OBESITY DRUGS IN THE PIPELINE

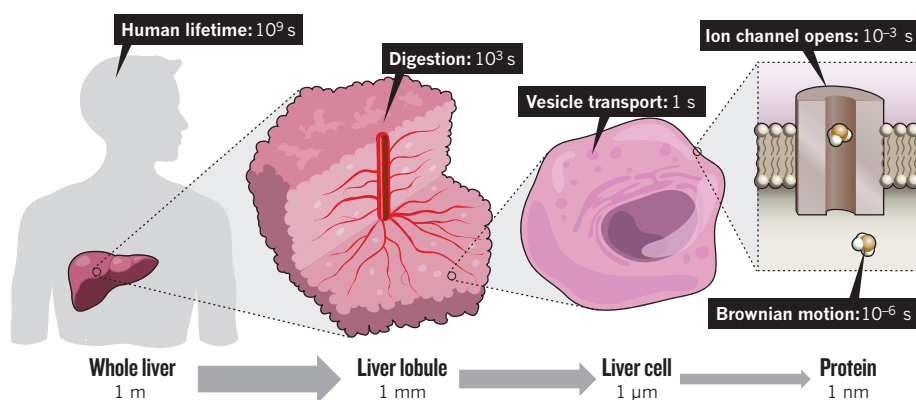
Drug	Developer	Target	Clinical trial status	Estimated approval date	Projected revenue
Victoza (liraglutide)	Novo Nordisk	GLP-1 receptor	Phase III	2015	\$780.4 million
Empatic (zonisamide/bupropion)	Orexigen Therapeutics	Dopamine and noradrenaline reuptake	Phase IIb	2017	\$443.0 million
Symlin (pramlintide)/metreleptin	Takeda/Amylin	Amylin receptor, leptin receptor	Phase IIb	2015	\$898.0 million
Velneperit	Shionogi & Co.	Neuropeptide Y/peptide YY receptors	Phase IIb	2014	—
Tesofensine	NeuroSearch	Dopamine, noradrenaline and serotonin reuptake	Phase IIb	—	—

SOURCE: BIOMEDTRACKER

SOURCE: GOOGLE FINANCE

## FROM MACRO TO NANO

Modelling human liver function means integrating data at scales from the whole organ down to molecules, and decades to microseconds.



## SYSTEMS BIOLOGY

# Germans cook up liver project

*Biologists join physicists in a bid to map the workings of the human organ at all scales.*

BY ALISON ABBOTT

**S**ystems biology — the holistic, interdisciplinary approach to modelling the processes of life — has set itself an ambitious new goal.

Launched in Dresden last week, the Virtual Liver Network is a German collaboration between biologists and theoretical physicists to model the functioning human liver. The work could help to develop more effective medicines, for example, because the metabolism of drugs in the liver has a profound impact on their efficacy and toxicity; or help in the understanding of liver disease. All foreign molecules are taken up into liver cells to be metabolized in preparation for excretion from the body.

Although some models of molecular pathways in liver cells can already predict how a drug may break down to become active or produce a toxic chemical, the biological consequences of this can only be predicted with a model of the liver's entire interacting system of cells and tissues. Similarly, a model of the entire liver will reveal much more about liver disease than a model of molecular interactions alone will.

The central challenge of the project lies in developing mathematical models that can unite data about processes that operate on vastly different temporal and spatial scales (see 'From macro to nano'). If successful, the network will integrate models of subcellular molecular signalling pathways with models of how a whole

cell works, eventually building up a model of the entire organ that will be available to drug developers and other researchers.

Success will also depend on meeting an equally tough sociological challenge — getting 250 scientists in 69 research groups around Germany to work towards a common goal. "The network is very demanding and its principal investigators all have many different pressures to attend to," says Ursula Kummer, a modeller from the University of Heidelberg. "But we are all very excited about the challenge of multiscale modelling, and that's what motivates us."

The German federal research ministry has provided €43 million (US\$57 million) to support the network for five years. It expands on the €36 million HepatoSys systems-biology programme, which between 2004 and 2009 worked towards modelling the hepatocyte, the most abundant cell in the liver.

The country's research community was hostile to HepatoSys at first, resenting the programme's top-down orchestration by the government's ministry of research. Many scientists also initially scorned the ministry's decision to use freshly dissected hepatocytes, rather than an off-the-shelf cultured liver cell line that would have been more convenient to use.

It took years for HepatoSys scientists to develop protocols that ensured each batch of new cells

would reach a standard quality in all research labs, but scientists have come round to the benefits. "We've found in fact that cell lines in culture don't behave much like real liver cells," says molecular biologist Ursula Klingmüller from the German Cancer Research Center in Heidelberg.

And Marino Zerial, a director at the Max Planck Institute of Molecular Cell Biology and Genetics in Dresden, points out that hepatocytes offer "a great system for experimenters" because they naturally take up foreign molecules. Unlike most other cell types, they readily incorporate RNAi — RNA interference molecules, which target individual genes in cells for suppression or silencing.

HepatoSys involved 47 researchers, and some collaborating biologists and physicists formed intense bonds. Zerial and Yannis Kalaidzidis, his colleague at the institute, describe themselves as "almost inseparable".

Few papers have emerged from the six-year programme, but funders consider it a success. "We needed that amount of time to learn to speak each others' languages and to work to a common goal," says Gisela Miczka, who administers the Virtual Liver Network for the research ministry. "Plenty of papers from HepatoSys will start coming out in the next two years." With other colleagues in Dresden, Zerial and Kalaidzidis took five years to develop a model that describes how nutrients or signalling molecules are transported into the cell. They are only now preparing to submit a manuscript on the findings.

The Virtual Liver Network will continue to generate data and develop spatio-temporal models of biological events on the cellular scale, but will also try to generate a new theoretical framework for systems analysis at every scale of liver function. The multiscale modelling will rely on standard tools including differential equations and stochastic statistical methods, says Kalaidzidis, but it will also require more elaborate mathematics to bridge data sets at different scales. "There is no general recipe for moving between scales," he says, so the researchers must work out the essential features at each scale before applying them at the next level.

This is not the first attempt to model a human organ. Physiologists around the world have been working on a computational model of the heart for a few years, and this effort set the stage for a European Union programme on modelling different organs — the Virtual Physiological Human Network of Excellence, launched in 2008 — on multiple scales. "But we haven't yet really bridged the scales much," says Peter Hunter, director of the Bioengineering Institute at the University of Auckland in New Zealand, who was a major driving force behind the programme. The large size of the Virtual Liver Network and the basic, single-scale work that has already been done, means the liver effort "is well-placed to make breakthroughs". ■

**NATURE.COM**  
See Nature's  
systems biology  
special at:  
[go.nature.com/lgerjv](http://go.nature.com/lgerjv)

## BIOINFORMATICS

# The rise of the genome bloggers

*Hobbyists add depth to ancestry trawls.*

BY EWEN CALLAWAY

Hours after Joseph Pickrell put his genome on the internet, an anonymous blogger took the data and concluded that he came from Ashkenazi Jewish stock. Pickrell, a genetics graduate student at the University of Chicago, Illinois, was sceptical about the claim. But after talking to relatives, he discovered that he had a Jewish great-grandfather who had moved to the United States from Poland at the turn of the nineteenth century. "It was a part of my ancestry I was totally unaware of," he says.

The blogger, who writes under the pseudonym Dienes Pontikos at <http://dodecad.blogspot.com>, had commandeered Pickrell's DNA as part of the Dodecad Ancestry Project, an ambitious project in which cutting-edge genomic analysis meets Web 2.0. Pontikos analyses genetic data submitted by followers

of his blog to reconstruct personal ancestry and human population history — and reports his findings online. He is part of a small but growing group of 'genome bloggers', a mix of professional scientists and hobbyists proving that widely available tools for computational biology could enable recreational bioinformaticians to make new discoveries.

"They are not amateurs. They are far from being amateurs," says Doron Behar, a population geneticist at Rambam Health Care Campus in Haifa, Israel, who studies human history. "I cannot stress enough the level of appreciation I have for their efforts."

Pontikos has so far analysed several hundred thousand single-letter DNA variations from more than 2,200 individuals. That includes more than 200 submitted to him by readers of his blog, who had had their genomes analysed by genetics testing firms such as 23AndMe, based in Mountain View, California, with the remainder coming from publicly available datasets. The

readers volunteering their genomes (identities stay private) are mostly keen to delve into their own ancestry. But Pontikos, who is from Greece and describes himself as an "anthropology diletante", is more interested in unfurling the history of populations that tend to be overlooked by human-population geneticists. For instance, his analysis of genomes from people living in northern Eurasia reveals a genetic connection between populations in northern Finland and central Siberia (see 'Meet the ancestors').

David Wesolowski, a 31-year-old Australian who runs the Eurogenes ancestry project (<http://bga101.blogspot.com>), also focuses on understudied populations. "It's a response, in a way, to the lack of formal work that's been

**"They are not amateurs. They are far from being amateurs."**

done in certain areas, so we're doing it ourselves," he says. Wesolowski and a colleague have drilled into the population history of people living in Iran and eastern Turkey who identify as

descendants of ancient Assyrians, and who sent their DNA for analysis. Preliminary findings suggest their ancestors may have once mixed with local Jewish populations, and Wesolowski plans to submit these results to a peer-reviewed journal.

But Pontikos sees little point in formally publishing his findings. "I can bypass them entirely, and have the entire world review what I write," he wrote in an e-mail. Indeed, comments on his blog — "could you please provide the eigenvalues for the principal component analysis", for instance — read like the niggling recommendations of a manuscript reviewer.

Pickrell notes that Dodecad and Eurogenes use cutting-edge techniques and open-source software developed by geneticists studying population history. The methods — which involve modelling past mixing between populations and distilling vast quantities of genotype data — still stir debate in the peer-reviewed literature because they can be difficult to interpret unambiguously, says John Novembre, a population geneticist at the University of California, Los Angeles. Behar, whose data on Jewish ancestry have been used by both projects, cautions that the techniques are more robust when applied to the history of an ethnic group, rather than the ancestry of an individual.

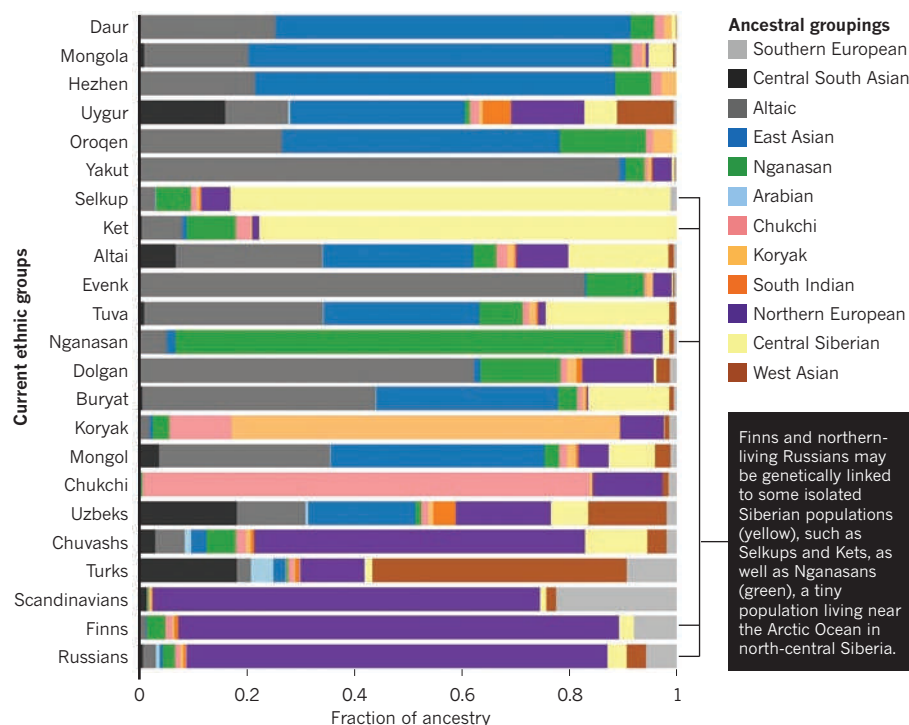
In response to concerns about the genetic privacy of those offering their genomes for analysis, "I don't think this is too worrisome," says Hank Greely, director of the Center for Law and the Biosciences at Stanford University in California. Both projects provide adequate privacy protection, he says, although they both could do a slightly better job at disclosing the risk of a release.

Although people may be happy to part with their genomes to

**NATURE.COM**  
The human genome at ten:  
[www.nature.com/humangenome](http://www.nature.com/humangenome)

## MEET THE ANCESTORS

An analysis of data from the Dodecad Ancestry Project highlights genetic links between different ethnic groups.





learn more about their ancestry, the genetic and trait data needed for biomedical applications are much harder, if not impossible, for amateurs to come by. Public repositories, such as the US National Institutes of Health's database of Genotypes and Phenotypes, tightly restrict access.

One effort to change that is the Personal Genome Project, which is spearheaded by George Church, a geneticist at Harvard Medical

School in Boston. The project aims to make the complete genome sequences and traits of 100,000 people freely available to anyone, with no strings attached. So far it has enrolled 1,000 participants and published near-complete genomes for 10 of them. Pickrell and 11 other scientists and genomics experts also added to the trove of freely available genomic data recently when they released their genetic data as part of a project called Genomes Unzipped.

Church argues that better access to high-quality data could help this kind of informal bioinformatics to flourish, enabling computer-savvy people to make important contributions to genomics, just as they have with online businesses such as Facebook. "It didn't take that much training to become a social-networking entrepreneur. You just had to be a good coder," he says. With bioinformatics, "I think we're in a similar position." ■

## INFORMATION TECHNOLOGY

# Developers call for handy lab aids

*Macmillan hopes to partner with scientists to turn software into commercial products.*

BY DECLAN BUTLER

It's an unfortunate truism that scientists often have better software tools for managing their music or family photos than they do for tracking their experiments and data. Major software companies tend to focus on much larger consumer and business markets, and offer little software for researchers.

As a result, others are rushing to occupy the niche, among them Digital Science ([www.digital-science.com](http://www.digital-science.com)), launched last week by *Nature's* parent company, Macmillan Publishers. Digital Science's strategy is not only to develop its own products — independently and with partner companies — but also to tap into the innovation of researchers themselves. With few existing products available to satisfy their needs, a growing number of scientists have developed their own tools with which to better organize their research lives.

The company has launched an open call for researchers who have written promising software to submit proposals for turning it into a commercial product. Researchers understand their colleagues' needs best, but often "don't have resources to turn their software into a polished product", says Timo Hannay, Digital Science's managing director and former publishing director of *nature.com*. The company aims to partner with researchers, or their start-ups, to provide them with the financial, developer and business resources they need.

The goal is to offer researchers tools that are as intuitive and user-friendly as well-designed consumer software. The company will initially focus on text-mining software, metrics-based tools to help institutions and funders better assess the performance of their funding and researchers, and lab-management software to help to keep tabs on anything from experiments to reagents.

There is certainly a massive need for better software to increase productivity at every stage of the complex scientific workflow, says Alexander Griekspoor, who founded



Got any nifty apps for your lab?

the company Mekentosj, based in Aalsmeer, the Netherlands, which produces software for molecular-biology applications. Sriram Kosuri, a bioengineer at the Wyss Institute for Biologically Inspired Engineering at Harvard University in Boston, Massachusetts, agrees. "Every bench scientist I know has their own, and self-admitted non-optimal way to organize their information," he says. Kosuri is a founder of OpenWetWare ([openwetware.org](http://openwetware.org)), a wiki for sharing lab protocols and data between biology groups worldwide. "I think there is an increasing willingness to pay for useful tools," he says.

Scientists struggling to organize thousands of article PDFs strewn across their hard drives have already embraced services such as Griekspoor's Papers ([mekentosj.com/papers](http://mekentosj.com/papers)), and London-based Mendeley ([www.mendeley.com](http://www.mendeley.com)), which bring the simplicity of iTunes to managing papers. Mendeley also offers social-networking facilities and other features.

Publishing giant Elsevier has recently entered the research-services market with its SciVal suite of metrics tools. It is one of several publishers that view providing institutions

with performance-measuring applications as a strategic move, says David Bousfield, the London-based vice-president and lead analyst at Outsell, a publishing and information consultancy. Elsevier has also launched SciVerse, a platform for searching and sharing content from Elsevier's own databases and the web. The product provides programming interfaces that allow researchers to build their own customized applications on top of content from Elsevier and other data sources, such as government databases.

Kosuri thinks that Digital Science's strategy of refining software created by researchers makes sense. "Getting the most important sets of tools developed more extensively would be tremendous," he says. But Michael Eisen, a geneticist at the University of California, Berkeley, and co-founder of the Public Library of Science, notes that many researchers have little enthusiasm for devoting time to perfecting software they have written. "I expect few would be interested," he says.

"We're not trying to convince anyone to create a commercial product if the desire isn't already there," Hannay counters. "We're trying to tap into the small but significant proportion of researchers who have identified an unmet need and are trying to do something about it."

Freely available code isn't an ideal alternative — most open-source software for scientists "sucks", says Eisen, although "a lot of it is really good and essential to what we do". But many lab-management tasks can be done using generic consumer open-source tools, he says. "The people I know mostly use wikis to keep track of stuff in the lab: they're free, flexible, and easy to set up and use."

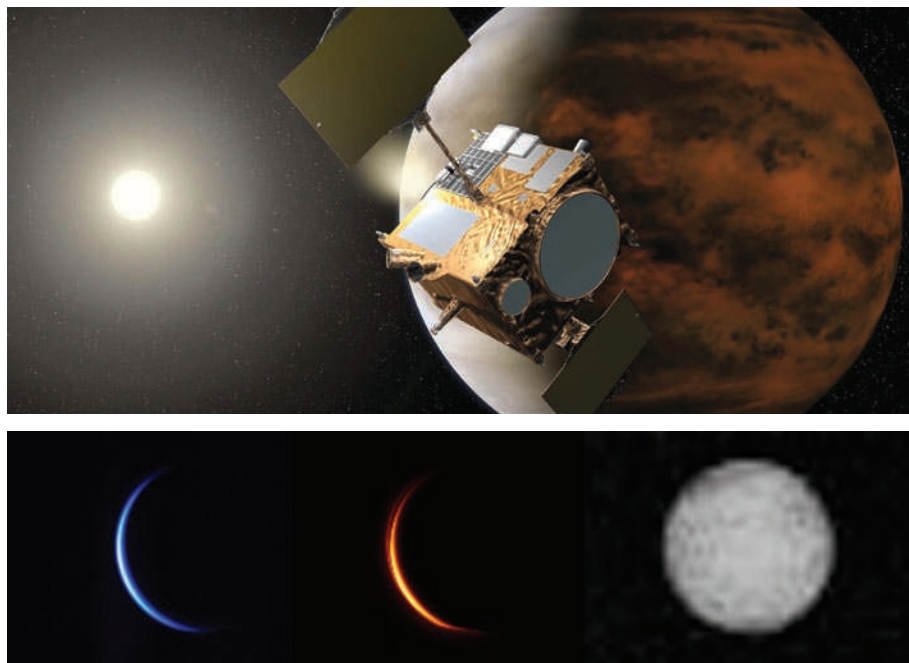
Eisen also warns against thinking of software as a panacea. "Everyone has aspirations to be better organized in the lab," he says, "they think there's a magic piece of software out there that will solve all their problems for them, but then they realize the problem is really that they're disorganized." ■

## ASTRONOMY

# Venus miss is a setback for Japanese programme

*Akatsuki mission on hold for six years before next attempt to approach planet.*

JAXA



Top: the craft was set to monitor Venus's atmosphere. Bottom: images taken as Akatsuki sped away.

BY DAVID CYRANOSKI

Few events can be as gut-wrenching for a planetary scientist as a multi-million-dollar spacecraft going silent while executing a crucial manoeuvre. Loss of signal at such times usually spells disaster, and the spacecraft may never be heard from again.

Researchers and engineers working with Japan's Akatsuki spacecraft were spared that worst-case scenario on 6 December. Although Akatsuki failed to make contact for more than an hour after the scheduled engine burn that was to place it in orbit around Venus, it did eventually call home. But the news was not promising. Not only had Akatsuki been tumbling out of control for a period of time, it had failed to enter orbit. It will now have to circle the Sun for six years before it gets a second chance.

The failure derails an ambitious programme of research into Venus's atmosphere, and marks the third time that the Japan Aerospace Exploration Agency (JAXA) has battled mechanical problems on a mission to another Solar System body. In 1998, a faulty valve caused a loss of fuel on JAXA's Nozomi spacecraft, which ultimately

prevented it from orbiting Mars. And the Hayabusa probe, which returned a minute quantity of asteroidal material to Earth this year, experienced a variety of near-fatal problems.

At a press conference on 10 December, officials reported that Akatsuki's engines fired for less than 3 minutes, far short of the 9 minutes and 20 seconds required to slip into orbit. "We are trying our best to get rid of any preconceived notions and figure out what happened," a project team member told *Nature*.

Akatsuki was to scour Venus with an infrared camera for evidence of volcanic activity, study lightning in the atmosphere and investigate the dense cloud layer that hides the planet's surface from view. Its planned equatorial orbit — timed to match the 'super rotation' of Venus's atmosphere, which spins about 60 times faster than the planet beneath it — would have allowed it to follow the long-term evolution of features in the cloud layer. Such data would have complemented the global coverage of Venus Express, the European Space Agency (ESA) probe that has been orbiting Venus since 2006.

"This is very disappointing for all of us," says Håkan Svedhem, an ESA project scientist for Venus Express. "We had hoped to do

many things jointly."

Engineers will now pore over telemetry data from Akatsuki and conduct tests with backup hardware on Earth to try to identify the source of the failure. The spacecraft's fuel system is likely to get close scrutiny: Akatsuki uses the same two-fluid hydrazine–nitrogen tetroxide thruster as Nozomi, although the valve issue has been addressed.

A series of images taken as Akatsuki sped away from Venus shows that its cameras are working well and may yet be put to use if the spacecraft survives its unplanned detour. One hurdle faced by Nozomi on its second attempt to enter orbit — frozen fuel — will probably not affect Akatsuki, because of the probe's proximity to the sun. Solar radiation might, however, take a toll on the craft's sensitive instruments.

If Akatsuki does reach Venus in 2016, it might still be able to join forces with other probes. Last month, ESA agreed to extend the Venus Express mission to 2017 or 2018, although this would require modifications to the craft's orbit to save fuel. Meanwhile, NASA is planning SAGE, a Venus lander that could launch in 2016.

"Understanding Venus is important because it informs us about the evolution of the climate on Earth," says Sanjay Limaye, an atmospheric scientist at the University of Wisconsin–Madison and a co-investigator on Akatsuki. "Not going into orbit now does not translate into a diminished interest in Venus, as the questions do not go away," he says. ■

**Akatsuki failed to make contact for more than an hour after a scheduled engine burn.**

## CORRECTION

The News story 'Self-plagiarism case prompts calls for agencies to tighten rules' (*Nature* **468**, 745; 2010) stated that Reginald Smith had escaped censure for research misconduct for publishing duplicate papers. In fact, Smith was formally reprimanded for reuse of published materials and data in multiple publications, although separate allegations of data falsification and plagiarism were not upheld.



# GROWTHINDUSTRY

To learn the chemical language of plants, Ian Baldwin has built up a German research empire that engineers seeds — and a field station in the Utah wilderness to grow them.

BY ALISON ABBOTT

In late spring 1988, Ian Baldwin was driving through the dessicating heat of the Utah desert in his rickety old VW microbus. The young researcher, from the State University of New York (SUNY), Buffalo, was searching for a native species of the tobacco plant as well as a place to sleep for the night. When he pulled up at the Desert Inn Ranch, he encountered a different form of wildlife. A posse of ferocious dogs flew out of the gate, puncturing his car tyres with their teeth. Behind them was rancher Herb Fletcher, cradling a submachine gun.

Baldwin was terrified. But when Fletcher called the dogs off, Baldwin slipped, very cautiously, out of the bus. Fletcher smiled — “He had a wonderful smile,” recalls Baldwin — and invited him in. The scientist and the old rancher quickly bonded over their shared interest in natural history. It was the start of a firm friendship — and the opening of a new era in Baldwin’s research life, one that has helped propel him into a dominant position in the burgeoning field of chemical ecology, the study of the chemical signals between plants and other organisms in the environment.

Rooted and unable to flee, plants have evolved many ingenious ways of repulsing their enemies, from generating noxious chemicals in their leaves to emitting complex, volatile bouquets to attract predators that will pick off the plant’s attackers<sup>1</sup>. It is a highly sophisticated chemical language undetectable by the human nose and largely undeciphered by science. But if and when it can be understood, it might

open the way to modifying plants’ signals to give them stronger protection, or to developing environmentally friendly mimics of natural signals as alternatives to herbicides.

In his efforts to understand this language, Baldwin has embarked on a project unique in its ambition and scale, carried out along what he calls “the longest lab corridor in the world”. Working in Jena, Germany, where he is a director of the Max Planck Institute for Chemical Ecology, he and his team develop powerful genetic tools to systematically knock out, or knock down, genes involved in making the chemical signals. Then they observe the effects by growing the modified plants in the wild — 8,844 kilometres away, next to the Utah ranch. The fastest journey from Jena to the field station takes 27 hours. The researchers have little choice, however. In Germany, with its populist aversion to anything genetically modified, such trials cannot proceed.

## COCKTAILS AND CRAZINESS

High-profile papers roll out of Baldwin’s institute with regularity. One published in *Science* in August<sup>2</sup> showed that the plants, when nibbled by herbivorous insects, can change the ratio of isomers of some of their signalling molecules specifically to attract predators of the leaf-eaters. And although Baldwin deliberately keeps his distance from applications, the agricultural industry studies his results attentively. It wants to learn how plants, or mixtures of plants, might be persuaded to produce the best cocktails of

volatile emissions for their own defences.

“Ian Baldwin is like a madman,” says Ted Turlings, a plant scientist at the University of Neuchâtel in Switzerland, with some awe. “He doesn’t stop working, day and night, and he lets nothing get in his way. He sets up a field station in an area where no one would think of going, builds himself a complete molecular tool set for his tobacco species, and sets out with purpose to get permission to use transgenic plants.”

For Baldwin, though, the approach is the only way to learn how a particular plant has evolved to survive in the real, stressful world of harsh weather and hungry insects. “It seems to me it would be madder not to do it this way,” he says.

With his humorous and low-key manner, and customary jeans, plaid shirt and baseball cap, Baldwin, 52, shows no sign of frenzy. The son of academic historians in Baltimore, Maryland, he decided early on that the ivory tower was not for him. “My parents are medievalists and live in the eleventh century,” he says. “I wanted to know what the rest of the world did for a living.” During high school and college he worked in his spare time as a fish cleaner, landscaper, truck driver, an auto and tractor mechanic, a logger and tree climber, and even a maple-sugar producer. And some of these skills became useful in unexpected ways.

As an undergraduate majoring in chemistry and biology at Dartmouth College in Hanover, New Hampshire, he became an informal assistant to Jack Schultz, one of the earliest pioneers of chemical ecology. Schultz, who was studying





Ian Baldwin looks after *Nicotiana attenuata* plants at one end of the world's longest lab corridor, his greenhouse in Jena, Germany.

forest canopies but couldn't stand heights himself, was as attracted to Baldwin's tree-climbing skills as to his precocious ability in the chemistry lab. In a joint experiment published in *Science* in 1983<sup>3</sup>, the pair claimed that chemicals from leaves that had been ripped to mimic insect damage could travel through the air to neighbouring plants and change their biochemistry in a way that wards off further insect attack. Their 'talking trees' notion was dismissed by many plant scientists as a fanciful over-interpretation of results.

## BALDWIN'S TREE-CLIMBING SKILLS IMPRESSED AS MUCH AS HIS PRECOCIOUS ABILITY IN THE CHEMISTRY LAB.

Burned by the reaction, Baldwin decided to play it safe for his PhD studies at Cornell University in Ithaca, New York, and instead researched the more mainstream internal signalling pathways within plants. The favoured plant model in the Cornell lab was *Nicotiana sylvestris*, a species of tobacco native to Peru. But when it came to extending this work into the field, Baldwin, by now on the tenure track at SUNY, decided to switch to a similar species native to the United States. His hunt for *Nicotiana attenuata* in Utah led to the fateful encounter with Fletcher.

Baldwin and his family spent the next seven summers with Fletcher, who regaled them with stories about the area's violent recent history. Fletcher had been brandishing a weapon at his first meeting with Baldwin because a neighbouring family of polygamists had been attacking him and his land. A mobster friend eventually put a stop to it.

If, during those summers, Baldwin learnt new things about how humans defend themselves, he learnt even more about plant defences. Fletcher would drive Baldwin around the 1,300-square-kilometre property to point out small clumps of *N. attenuata* he had spotted while out ranching, and Baldwin would use them in experiments to find out, for example, how the plants activated chemical defences against herbivores.

### THE HUNT FOR HOTSPOTS

After a brush fire in 1992, Baldwin discovered that the seeds of *N. attenuata* germinate only when activated by components of wood smoke penetrating the soil around seeds. Then they suddenly flourish in the temporarily nutrient-rich, herbivore-free, post-fire environment. He learnt to locate natural populations of *N. attenuata* more efficiently by chasing lightning strikes — "or simply phoning the fire department and asking them where they spent money". One of his earliest series of experiments was designed to understand the costs and benefits to the plant of one particular defence mechanism — producing nicotine in its roots and then pumping the toxin up into its leaves — in the field.

Plant biology had been transformed in 1990 with the discovery that the hormone jasmonic acid could induce volatile signalling. Baldwin immediately set out to see if it could also induce nicotine production, and found that it could. So he synthesized an artificial version of the hormone that he could inject into the soil around plant roots. His complex set of experiments involved four populations of *N. attenuata*, each with more than 1,000 plants, spread over a 100-kilometre loop. He found that among plants pretreated with jasmonate to artificially induce a level of nicotine defence, those that were not attacked produced less seed than those that subsequently were attacked, but not ravaged, by herbivores<sup>4</sup>. With this work, he proved a point about plant defence that had previously only been assumed — that using defences only when needed is an evolutionary advantage, because it maximizes benefits and minimizes cost.

More than a decade after his ill-received 'talking trees' *Science* paper, Baldwin was itching to return to the theme. The work had been vindicated by the early 1990s when several scientific groups working on crop plants had established that such plant volatiles not only exist, but stimulate responses from other species — pathogens, herbivores, herbivore-eating carnivores, and perhaps other plants as well. But most of the work had been done under laboratory conditions.

Baldwin wanted to understand plant biology in the real world. He particularly wanted to



explore the volatile chemicals that plants exude, using genetic manipulation to take apart the machinery involved in making them. He was convinced that the desert-dwelling *N. attenuata* would make an ideal model because it has evolved such an array of mechanisms to survive severe environmental stresses including fire, herbivores and drought.

What wasn't yet available for *N. attenuata*, however, was the toolkit necessary for genetic engineering. One was already being assembled for *Arabidopsis* — the favoured study subject of lab-based plant biologists, but one that Baldwin regarded as “a boring weed of no use to ecological evolution research”. He wanted to have the same for his *N. attenuata* — but he knew it would be expensive and take many years.

That didn't discourage the Max Planck Society in Munich, which recruited him in 1995. Following the reunification of Germany in 1990, the society was obliged to extend its network of research institutes into former East Germany, and it took the opportunity to add exciting areas of research and recruit more foreign directors. The society wholly bought into Baldwin's vision, and made him a founding director of the Max Planck Institute for Chemical Ecology in Jena. All Max Planck directors are given generous, guaranteed funding and time to develop long-term projects without having to apply for grants, making Baldwin's dream possible.

Baldwin now has a formidable institute employing more than 50 researchers, students and technicians to develop the chemical, as well as genetic tools to mimic or block signalling pathways.

Not even the might of the Max Planck Society could help him realize one part of his original vision — to stage field trials of genetically

## SUNGLASSES AND IPODS ARE FORBIDDEN AT THE FIELD STATION TO ENSURE PEOPLE AVOID THE RATTLESNAKES.

modified plants native to Germany. “Even if you do get approval, the German government now requires the GPS positioning of all field trials with transformed plants to be posted on the web — so every single trial is destroyed by activists,” he says.

In the United States, the work is arduous but doable. The team in Jena engineers seeds and sends them to the US Animal and Plant Health Inspection Service in Rockville, Maryland, where they are inspected and sent on to the field station in Utah. This season, Baldwin's team planted out 4,000 seedlings for around 36 studies on topics ranging from plant–pollinator interactions to how plants allow their roots to be colonised by microbes. “This is really a lot of backbreaking work,” says Baldwin.



In Utah, 8,844 kilometres from their German lab, Baldwin and his team study genetically modified plants.

Life at the field station is tough in other ways. Brush fires spread “faster than you can run”, says Baldwin. In 2005, a fire spotted on the horizon sent a dozen or so scientists running for their lives. Aware that their station has an explosive tank of propane fuel, they tore away in vans. (The fire shifted direction before reaching the station.) Baldwin admits to being “tyrannical” about safety. In a region shared with deadly animals such as the sidewinder rattlesnake, he forbids iPods and sunglasses — “people have to be able to hear and see snakes” — and no one is allowed to wander around the desert alone. It is an hour's rough drive to the nearest hospital. Everyone must learn how to change a tyre on the vans.

### LIFE IN THE WILD

Nature can add to their troubles in other ways. This year, for example, an elaborate series of experiments designed to study how the empoasca leaf hopper, a herbivore, recognizes and interacts with its host plant went to waste because the leaf hopper didn't show up.

But results continue to flow. One of Baldwin's most colourful papers, published in February this year<sup>5</sup>, probes the dilemma of plants that need to attract pollinators while remaining inconspicuous to herbivores. *Nicotiana attenuata* normally flowers at night, emitting the volatile benzyl acetone to attract hawk-moths. Unfortunately, hawk-moth larvae are also herbivores, and the moths often leave their eggs on the leaves as they pollinate. When the plants become infested, the team found, they shut down production of benzyl acetone and open their flowers at dawn when the moths are gone. They are then pollinated by hummingbirds. Using a series of genetically modified strains, Baldwin's team showed how the oral secretions

from the munching hawk-moth larvae trigger the dramatic switch in flowering time.

Baldwin's research has inspired attempts to develop new crop strains that could be practical for farmers in poorer countries who can't afford lots of pesticides and herbicides, says John Pickett, director of Rothamsted Research in Harpenden, UK, a historic agricultural research centre. The centre is working with some agricultural companies on field trials in the United States and United Kingdom of crops genetically modified to amplify chemical signals that plants make when they are under attack. Details are currently confidential, he says, but “we'll have big announcements in the next year or two”.

Back in Jena, at the end of the 2010 season, Baldwin is planning next year's experiments with genetically altered strains designed to have elevated or suppressed emissions of volatile signals, which he labels as the ‘screamers’ or the ‘mute’. He will investigate how single screamers planted among a colony of mutes, for example, might affect herbivore or predator behaviour.

Right now, Baldwin is 8,844 kilometres away from the next experiment at his barren, hostile field site. Human relations in the region are now a lot tamer. Relations between plants and insects, though, are as wild as ever. ■

**Alison Abbott** is Nature's senior European correspondent.

1. Baldwin, I. T. *Proc. Natl Acad. Sci. USA* **95**, 8113–8118 (1998).
2. Allmann, S. & Baldwin, I. T. *Science* **329**, 1075–1078 (2010).
3. Baldwin, I. T. & Schultz, J. C. *Science* **221**, 277–279 (1983).
4. Farmer, E. E. & Ryan, C. A. *Proc. Natl Acad. Sci. USA* **87**, 7713–7716 (1990).
5. Kessler, D., Diezel, C. & Baldwin, I. T. *Curr. Biol.* **20**, 237–242 (2010).

# COMMENT

**TECHNOLOGY** Lessons for future of the Internet in history of communications **p.892**

**DRUGS** Opium dominates show on legal and illegal highs **p.896**

**ECONOMICS** Beware, politicians will exploit any indicator **p.897**

**OBITUARY** Allan Sandage, who measured the Universe's expansion, remembered **p.898**



## Build life to understand it

Biologists and engineers should work together: synthetic biology reveals how organisms develop and function, argue **Michael Elowitz** and **Wendell A. Lim**.

**T**his year's publicity about Craig Venter 'creating' life<sup>1</sup>, and this week's report on the promise and perils of synthetic biology from US President Barack Obama's commission on bioethics, threaten to obscure the most important impact of this field. Synthetic biology is redefining the discipline of biology and helping people reach a deeper understanding of how life works.

Conventionally, biologists have sought to understand life as it exists. Increasingly, however, from stem-cell reprogramming<sup>2</sup> to microbial factories<sup>3</sup>, researchers are describing what is and exploring what could be. An analogous shift occurred in physics and

chemistry, especially in the nineteenth century. Like biology, these fields once focused on explaining observed natural processes or material, such as planetary motion or 'organic' molecules. Now they study physical and chemical principles that govern what can or cannot be, in natural and artificial systems, such as semiconductors and synthetic organic molecules<sup>4</sup>.

The expansion of biology from a discipline that focuses on natural organisms to one that includes potential organisms (see 'Beyond the natural') will have three long-term effects. First, it will enlarge the community of biologists to include researchers with different

assumptions and goals, such as engineers. Second, it will alter the way in which scientists address the fundamental problem of how biological systems work. Integrating reverse and forward engineering approaches will free biologists to uncover fundamental principles that explain, unify and extrapolate beyond mechanisms observed in specific model systems. Third, it will provide a new conceptual basis for teaching biology — one founded on stimulating inquiry from students as to how biological components and modules could be used to implement complex functions.

Although traditional disciplinary boundaries are dissolving, the cultural differences ►

YELLOW BY NATHAN SAWAYA/BRICKARTIST.COM



► between scientists and engineers remain strong. For biologists, genetic modification is a tool to understand natural systems, not an end in itself. Thus, making biological systems ‘engineerable’ — a goal of engineers in the field of synthetic biology — can seem pointless. Many biologists wonder why engineers fail to appreciate the intricate, beautiful and sophisticated designs that occur naturally. Engineers are often equally perplexed by biologists. Why are they so obsessed about the details of one particular system? Why don’t they appreciate the value of replacing a complex and idiosyncratic system with a simpler, more modular and more predictable alternative? These misunderstandings can make for fascinating conversations, but they can also prevent mutually beneficial synergies.

Biologists and engineers need to appreciate the complementarity of their approaches. Below the surface, these two communities have common interests and goals that can, and must, be addressed from both directions — forward and reverse engineering.

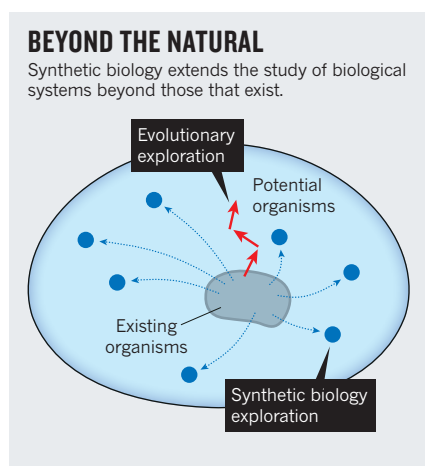
### SAME CHALLENGES

Traditional biologists seek to reverse engineer natural biological systems — to understand how their molecular circuitry, composed of interacting genes and proteins, gives rise to observed behaviour. Synthetic biologists seem to do the opposite. They forward engineer new behaviour using well-understood genetic components and as simple a design as possible. Both communities face the same daunting challenge: how to relate the architecture of a gene circuit to its behaviour in a cell or tissue.

Synthetic circuits can provide insights into natural circuit-design principles that would be difficult or impossible to obtain using conventional perturbations of natural systems alone. Consider signalling. Biologists have discovered that a handful of canonical pathways are used repeatedly across species, tissues and stages of development. What is it about this set of pathways that makes it sufficient for the development and physiological function of a complex organism? To address this question requires an understanding of what each pathway can do. Synthetic biologists can systematically engineer a diverse range of signalling-pathway architectures and analyse them in relative isolation from any particular set of downstream processes<sup>5,6</sup>. These architectures may include natural, as well as new, configurations. The results could provide a higher-level view of signalling in which one could associate each pathway and architecture with a specific functional repertoire, instead of thinking about them primarily in terms of their molecular interactions.

A second example of where synthetic biology can provide a complementary approach is metabolic networks — one of the most active frontiers in the field. Biology has

conventionally focused on understanding the metabolic pathways in particular organisms. Synthetic biology enables researchers to consider what types of metabolic networks are possible by combining enzymes from all species. Such work has focused on engineering novel metabolic pathways that produce specific molecules for medicine and industry. These efforts can also address fundamental biological questions. For example, what trade-offs exist between metabolic efficiency and flexibility? Are there fundamental principles for how cells set up their metabolic economy and synthesize and distribute key chemical precursors<sup>7,8</sup>? These questions could be important for understanding the diversity of metabolic networks in natural microbes as well as in biomedically important systems such as cancer, in which cells alter their metabolism<sup>9</sup>.



Synthetic-biology approaches may also provide insights in developmental biology. They could be used to tackle fundamental questions of what types of multicellular patterning processes are possible, and what types of circuits — combining signalling, regulation, differentiation and morphological change — would be sufficient to program the formation of organisms.

Using well-characterized signalling pathways, transcription factors and regulators of cell morphology and division, it should become possible to explore a range of natural and non-natural developmental circuit architectures. This would start with very simple patterns that could be generated in relative isolation from other developmental processes in the simplest systems. Eventually, synthetic developmental systems should yield a deeper understanding of morphological programming, provide insights into natural developmental systems and possibly enable applications in tissue engineering.

The convergence of engineering and biology could bring exciting new ways of teaching biology. Conventional biology, focused on understanding the structure,

mechanism and origins of extant beings, tends to involve memorizing nomenclature and facts. In some cases, this approach can obscure unifying principles and concepts.

Instead, teachers could start by challenging students with the question: ‘how might you build a biological system that performs a particular function?’ Students could be asked to deduce underlying design principles — for example, to identify the general types of circuit modules necessary or sufficient to implement a given behaviour in cells. Students thus equipped with organizing concepts could better navigate the sea of confusing nomenclature in natural living systems. Inconsistencies between idealized designs and actual examples would raise important questions about assumed functions, and about constraints inherent to the evolutionary process. In physics and engineering, this kind of approach is commonplace and can be effective at engaging and motivating students.

Such concepts could be introduced to teenage students who are just starting to think more deeply about the mechanisms underlying plants and animals. Requiring theory, computation and experiment would better equip students for multidisciplinary research. It would also expose them, at an earlier stage, to the conceptual and creative aspects of the scientific process, potentially attracting a broader range of people to biology.

Many technical and fundamental obstacles remain before the design and construction of synthetic biological systems can become routine. And as discussed in the commission report, the societal challenges may be equally formidable. Bringing together the energies and expertise of diverse communities that think about biological problems in different terms is a good first step towards taking full advantage of the many opportunities that lie ahead. ■

**Michael Elowitz** is in the Division of Biology, California Institute of Technology, Pasadena, California 91125, USA. **Wendell A. Lim** is in the Department of Cellular and Molecular Pharmacology, University of California, San Francisco, California 94158, USA.  
e-mails: melowitz@caltech.edu; lim@cmp.ucsf.edu

- Gibson, D. G. et al. *Science* **329**, 52–56 (2010).
- Takahashi, K. & Yamanaka, S. *Cell* **126**, 663–676 (2006).
- Carothers, J. M., Goler, J. A. & Keasling, J. D. *Curr. Opin. Biotechnol.* **20**, 498–503 (2009).
- Yeh, B. J. & Lim, W. A. *Nature Chem. Biol.* **3**, 521–525 (2009).
- Bashor, C. J., Horwitz, A. A., Peisajovich, S. G. & Lim, W. A. *Annu. Rev. Biophys.* **39**, 515–537 (2010).
- Sprinzak, D. & Elowitz, M. B. *Nature* **438**, 443–448 (2005).
- Bar-Even, A., Noor, E., Lewis, N. E. & Milo, R. *Proc. Natl Acad. Sci. USA* **107**, 8889–8894 (2010).
- Noor, E., Eden, E., Milo, R. & Alon, U. *Mol. Cell* **39**, 809–820 (2010).
- Hsu, P. P. & Sabatini, D. M. *Cell* **134**, 703–707 (2008).



S. KAZLOWSKI/SCIENCE FACTION/CORBIS

The polar bear (possible hybrid pictured) is one of several species vulnerable to hybridization.

# The Arctic melting pot

Hybridization in polar species could hit biodiversity hard, say **Brendan Kelly, Andrew Whiteley and David Tallmon.**

In 2006, a white bear with patches of brown fur was shot by hunters in the Arctic. DNA tests confirmed what many suspected — it was a hybrid of a polar bear and a grizzly. A media frenzy quoted biologists as saying that although they knew in theory such cross-breeding could happen, they didn't expect to see it in the wild. In 2010, another hybrid was killed by a hunter in the western Canadian Arctic. This time, the animal was a second-generation cross — its mother was a hybrid and its father a grizzly. More cases are probably out there.

Biologists should not be surprised. There have been hints of Arctic hybrids before. In the late 1980s, a whale thought to be a narwhal-beluga mix was found in west Greenland. In 2009, an apparent bowhead-right-whale hybrid was photographed in the Bering Sea, between Alaska and Russia. Dall's porpoises are known to be mating with harbour porpoises off the coast of British Columbia, and seal hybrids have been identified in museum specimens and in the wild.

These are just the first of many hybridizations that will threaten polar biodiversity. Rapidly melting Arctic sea ice imperils species through interbreeding as well as through habitat loss. As more isolated populations and species come into contact, they will mate, hybrids will form and rare species are likely to go extinct. As the genomes of species become mixed, adaptive gene combinations will be lost.

Researchers have little idea how much hybridization is occurring, let alone how it will affect populations. Plans must be

developed immediately to monitor the genetics of Arctic animals and to deal with hybrids before currently discrete populations merge and at-risk species are bred out of existence.

We have counted at least 34 possible hybridizations between discrete populations, species and genera of Arctic and near-Arctic marine mammals (see Supplementary Information). Of the 22 species involved, 14 are listed — or are candidates for listing — as endangered, threatened or of special concern by one or more nations. Twelve cases are of hybridization between different species — half involving crosses between what are normally classified as distinct genera. Twenty-two cases involve isolated populations at risk of intra-species mixing, nine of which are classified as distinct subspecies.

The Arctic Ocean is predicted to be ice-free in summer before the end of the century, removing a continent-sized barrier to interbreeding. Polar bears are spending more time in the same areas as grizzlies; seals and whales currently isolated by sea ice will soon be likely to share the same waters.

Not all cross-species matings will produce viable — or indeed any — offspring. The chance is enhanced in Arctic marine mammals, because their number of chromosomes has changed little over time. There is evidence of hybridization across species (such as between spotted and harbour seals) as well as across genera (such as harp and hooded seals).

Hybridization is not necessarily a bad thing. It has been an important source of evolutionary novelty. For example, a new species of chub originated in the Colorado River before

the presence of humans, from the hybridization of two other species. But hybridization driven by human activities tends to occur quickly and to reduce genomic and species diversity. When mallard ducks were introduced to New Zealand in the 1860s, they began mating with native grey ducks. Now few, if any, pure native populations remain.

Diversity loss may be minor if, say, North Pacific and North Atlantic minke whale subspecies interbreed in an Arctic with diminished ice. Other crosses will be more problematic. Interbreeding between the North Pacific right whale, of which there are probably fewer than 200, and the more numerous bowhead whale could quickly push the former to extinction. If polar bears survive climate change in secluded refuges — which is far from certain — interbreeding could be the final straw.

Cross-breeding might affect social and ecological interactions. The apparent narwhal-beluga hybrid discovered in Greenland had teeth combining qualities of each species, but lacked the narwhal's tusk — an important determinant of narwhal breeding success. Polar-grizzly hybrid bears in a German zoo exhibited behaviour associated with seal hunting, but not the strong swimming abilities of polar bears. First-generation crosses can have 'hybrid vigour', but later generations are likely to be less fit than their ancestors ('outbreeding depression').

The International Union for the Conservation of Nature should develop a comprehensive policy for managing hybrids, including determining when it is practical to prevent or limit hybridizations. Red wolf and coyote hybrids, for example, have been culled in the United States in the past decade to help preserve distinct species.

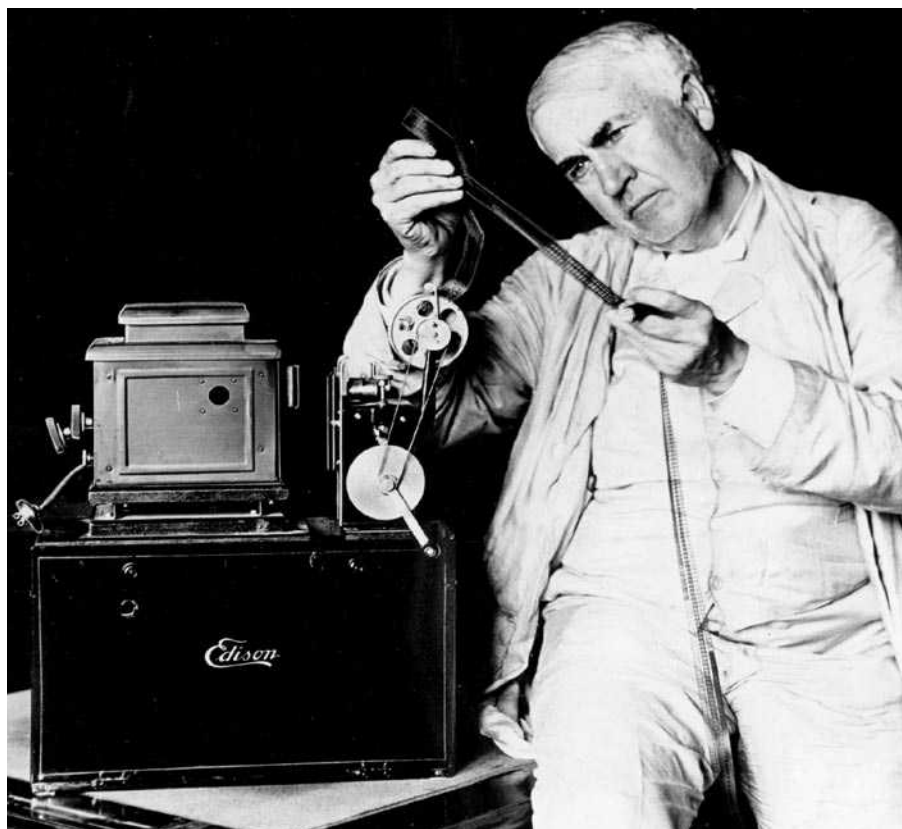
Researchers should combine models of sea-ice loss, oceanography and landscape genomics to predict when and where hybridization is most likely, and to monitor the genetics of at-risk populations. National and tribal governments should work together: some indigenous groups actively monitor the harvests of Arctic marine mammals, and they could collect genetic samples in remote areas. The rapid disappearance of sea ice leaves little time to lose. ■

**Brendan P. Kelly** is at the National Marine Mammal Laboratory National Oceanic and Atmospheric Administration, Juneau, Alaska 99801, USA. **Andrew Whiteley** is in the Department of Environmental Conservation, University of Massachusetts, Amherst, Massachusetts 01003, USA.

**David Tallmon** is in the Faculty of Biology and Marine Biology, University of Alaska Southeast, Juneau, Alaska 99801, USA. e-mail: [brendan.kelly@noaa.gov](mailto:brendan.kelly@noaa.gov)

Further reading and Supplementary Information accompanies this article at [go.nature.com/h4bksj](http://go.nature.com/h4bksj)





Thomas Edison invented and patented many early film technologies, allowing him to control the industry.

## INNOVATION

# Fighting monopolies

A history of communications technologies holds lessons for the Internet today, finds **Li Gong**.

Innovation and business interests do not always mesh. Thomas Edison, the inventor of the light bulb and the phonograph, almost suffocated the US film industry in the early 1900s by controlling all the crucial patents for film technology. His Motion Picture Patents Company (also called the Edison Trust) dictated film length, style, content, who could show films and at what price.

In 1934, New Jersey's Bell Laboratories, the birthplace of the semiconductor, suppressed development of its magnetic tape and the answering machine for six decades to protect the telephone business of its corporate parent, AT&T, who feared that the recording of conversations would deter people from using telephones. The arrival of fibre optics, mobile phones, faxes and speakerphones were similarly delayed.

In his groundbreaking book, *The Master Switch*, Columbia University law professor Tim Wu weaves together these and other examples to examine how disruptive technologies enter and develop within society. The new industries that emerge, he argues, progress in a cycle: companies grow to become empires, which close the field until the next wave of technology arrives to dismantle the existing order.

Wu covers the histories of radio, music, film, television and the Internet. All are littered with examples of vested interests that have thwarted competition and reduced innovation through commercial, political, legal and regulatory pressures. Drawing on their substantial war chests, large companies can lobby hard. For example, one telecommunications giant persuaded the state of

Texas in 1995 to pass a law requiring that any would-be companies must build phone lines that reach at least 60% of homes and businesses, a measure that shut out new competitors. Another tactic is to charge exorbitant rental fees for facilities owned by large companies.

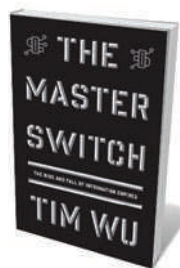
Monopolistic powers also restrict freedom of expression and civil liberty. For example, after deciding arbitrarily that films more than a few minutes in length were uninteresting, the Edison Trust refused to license longer feature films. Studios such as Paramount Pictures, Fox and Universal sprang up in rebellion. Hollywood grew partly as a result of its proximity to Mexico, where independent filmmakers could escape from injunctions and subpoenas coming from controlling interests on the US East Coast.

On taking over ownership of the film industry, these few studios soon applied their own censorship. Bowing to pressure from Catholic activists to uphold moral values on screen, the film-industry bosses in 1934 agreed to abide by a production code, known as the Hays Code. Named after campaigner William Hays, president of the Motion Picture Producers and Distributors of America, the set of rules specified what was considered obscene. For decades it restricted what the public could view.

## PHONE WARS

Democracy is also influenced by the narrow ownership of telecommunications. For example, in the US presidential election of 1876, news monopoly Associated Press supplied content for communications monopoly Western Union, which leaked rivals' confidential telegrams to its favoured candidates. In the past decade, the US government has been accused of secret wiretapping, made possible by the concentration of telecommunications in a few hands. Thus there is constant combat over the ownership of communication technologies — between open and closed, decentralized and centralized models.

The latest battlefield is the Internet. Wu stresses that, because it is so crucial to our society, we must prevent the monopolistic cycle that might close this diverse, distributed, decentralized and democratic system. Such attempts to gain broad control have so far failed — witness



**The Master Switch: The Rise and Fall of Information Empires**

TIM WU  
Knopf: 2010. 384 pp.  
\$27.95

➔ **NATURE.COM**  
Carl Zimmer muses  
on science and film:  
[go.nature.com/geo44i](http://go.nature.com/geo44i)



the ill-fated merger, now dissolved, between media giant Time Warner and Internet service provider AOL. But Wu is wary of the rise of 'closed' platform devices that restrict what programs can be used, such as Apple's Mac, iPad and iPhone, compared with open systems, such as the earlier Apple II. He quotes Tom Conlon writing online in *Popular Science*: "Once we replace the personal computer with a closed-platform device such as the iPad, we replace freedom, choice, and the free market with oppression, censorship, and monopoly."

Central in keeping the Internet open is the concept of 'network neutrality', which Wu has popularized: government and information carriers should place no restriction on where, when, what and how users access information. A requirement of net neutrality is that Internet service providers should not use price differentiation to fend off upstarts, to favour their collaborators, or to retain their monopolistic power in new or adjacent fields. But opinions are varied and examples to the contrary abound: the US cable firm Comcast allegedly levied additional fees for video traffic from companies that compete

with its cable business, for example. The British government has also recently announced support for a two-speed Internet.

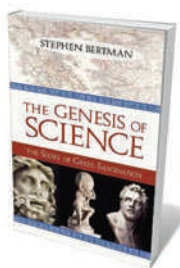
Wu believes that the antiquated competition laws

that focus on pricing to protect consumers are inadequate in the information industry, because collusion restricts choices but does not always inflate prices. Rather than legislation, he proposes a 'separations principle', whereby vital components of the information industry are entrusted to different institutions, both public and private. These bodies would apply checks and balances to ensure that control is not given to only a few players. Such an idea is attractive, yet will undoubtedly be difficult to put into practice because of vested interests.

*The Master Switch* offers powerful lessons from the past for the future of the Internet. Should we let it evolve along its natural trajectory, and risk it becoming temporarily controlled by monopolies until the next breakthrough? Or should we intervene to protect innovation and a free and open Internet at any expense? Perhaps, though, we don't have as much control as we think. Wu cites ancient Chinese wisdom from Luo Guanzhong: "An empire long united, must divide; an empire long divided, must unite. Thus it has ever been, and thus it will always be." ■

**Li Gong** is chief executive of Mozilla China.  
e-mail: lgong@mozilla.com

## Books in brief



### **The Genesis of Science: The Story of Greek Imagination**

Stephen Bertman PROMETHEUS BOOKS 304 pp. \$27 (2010)

The origins of science in ancient Greece are explored by classicist Stephen Bertman. He looks beyond the familiar names such as Euclid and Pythagoras to lesser-known figures, including the mapmaker Anaximander and alchemist Maria the Jewess, popularly known for inventing the eponymous bain-marie water bath and various pieces of chemical apparatus, including the still. Bertman argues that the Greeks owe their scientific success to their belief in an ordered Universe, the rules of which could be unpicked by the human mind.

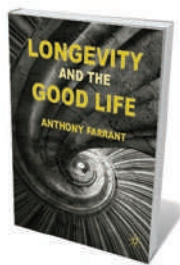


### **Hunger: The Biology and Politics of Starvation**

John R. Butterly and Jack Shepherd DARTMOUTH COLLEGE PRESS

360 pp. \$29.95 (2010)

One in seven of the world's population is short of food. Lack of political will is the main reason for not addressing hunger, explain medical scientist John Butterly and environmental scientist Jack Shepherd. As well as describing the biology of human nutrition and famine, they examine the political and historical factors that cause hunger and malnutrition to remain major health problems today despite advances in science and technology and the proliferation of humanitarian efforts.



### **Longevity and the Good Life**

Anthony Farrant PALGRAVE MACMILLAN 256 pp. \$85 (2010)

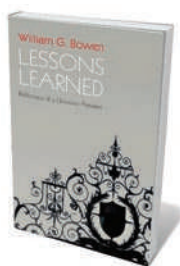
Living longer may not be such a good thing, cautions bioethicist Anthony Farrant. Although breakthroughs in medical biotechnology have the potential to extend our lives and make them healthier, he disputes the idea that immortality is desirable and cautions that the ready availability of such enhancements will diminish the value we put on reaching old age. Increasing longevity will challenge the fair distribution of resources, especially health care. Ultimately, he says, these pressures will undermine the idea that all people are fundamentally equal, and thus threaten the good life.



### **Man and Woman: An Inside Story**

Donald W. Pfaff OXFORD UNIVERSITY PRESS 232 pp. £15.99 (2010)

Gender differences have deep and tangled roots, according to neuroscientist Donald Pfaff. Although genetic and biological factors such as neuroanatomy contribute to this dichotomy, he argues, they do not dominate. Cultural influences, including experiences of stress throughout various stages of our lives, may be just as large and affect males and females in varied ways. Differences between the sexes, both physical and mental, result from a combination of genetics and environment that operates on many levels to influence behavioural mechanisms.



### **Lessons Learned: Reflections of a University President**

William G. Bowen PRINCETON UNIVERSITY PRESS 168 pp. \$24.95 (2010)

William Bowen reflects on the lessons he learned while he was president of Princeton University in New Jersey from 1972 to 1988, and president of the Andrew W. Mellon Foundation in New York from 1988 to 2006. He shares advice on fund-raising, hiring, managing faculty members and interacting with trustees. And he reveals his experience of shepherding the elite university through the civil-rights movement and the Vietnam War, a period during which he helped to expand the faculty, especially in the life sciences.





Eels are revered as gods by some cultures, including that of the Maori — as depicted in this wall mural in Canterbury, New Zealand.

## ECOLOGY

# The mystery of eels

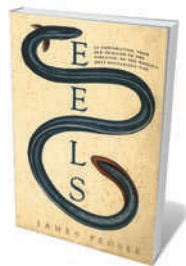
Kim Aarestrup is reminded of how little we know about these endangered fish.

Snake-like and nocturnal, eels are mysterious creatures. They spawn in remote and nutrient-poor places in the seas, and no human has ever seen one reproduce in the wild. Their rice-sized hatchlings embark on an odyssey of up to 6,000 kilometres to find fresh or brackish water, where they grow for decades — reaching weights of more than 20 kilograms — only to return to the sea, where they spawn, die and sink into the abyss.

Exploited as food for millennia owing to their abundance, taste and high energy content, eels cannot yet be cultured profitably. All traded eels are wild — and populations are plummeting. Species in temperate areas, including the American, Japanese and European eel, have become scarce, with populations dropping by more than 90% in the past four decades. European eels are now listed as critically endangered by the International Union for Conservation of Nature, a shocking development for a fish once found across all the accessible waters of Europe.

In *Eels*, naturalist James Prosek travels

and interviews leading scientists worldwide to examine the *Anguilla* genus. By broadening the perspective beyond Atlantic species, his book complements Tom Fort's marvellous *The Book of Eels* (HarperCollins, 2002). As well as describing the biology of the eel,



**Eels: An Exploration, from New Zealand to the Sargasso, of the World's Most Mysterious Fish**  
JAMES PROSEK  
HarperCollins: 2010.  
304 pp. \$25.99

Prosek considers the cultural and economic value we attach to it, interweaving historic vignettes from Aristotle's interest in the origins of European eels to Sigmund Freud's nineteenth-century search for their testes.

Prosek visits New Zealand, where the Maori revere the large endemic longfin eel *Anguilla dieffenbachii* as a religious symbol, which they believe can bark like a dog

and scream like a baby. Large road projects have been diverted to avoid areas populated by *taniwha*, or special guardian eels. Prosek goes to Japan, a nation that eats vast quantities of eel, making the traditional dish *kabayaki* a multimillion-dollar industry. He also visits the Micronesian island of Pohnpei, where *Anguilla marmorata* is sacred, believed to be the islanders' ancestor.

Restoration of eel populations will be difficult. Prosek lists contributors to their decline: loss of habitat, dams, fishing, introduction of parasites, pollutants and changes in ocean currents. These factors, and our lack of knowledge about key stages of the eel life cycle, make population management problematic.

The plight of temperate species has led to a surge of eel research in the past few years. Recent papers have described captures of Japanese eels that have spawned, showing that they do so in tropical ocean frontal zones, a mixing zone between warm and cold oceanic waters. Other research has revealed the diet of newly hatched eel larvae (called leptocephali)



## CONSERVATION

# Biodiversity as a bonus prize

Rare species and ecosystem services make uneasy bedfellows, discovers **Emma Marris**.



and suggested alternative larval migration routes for European eels other than the North Atlantic drift current. Prosek's book stops short of capturing these emerging results.

For American and European eels, monitoring the late stages of their life cycle in the Atlantic is the greatest challenge. Only by assessing survival rates can we focus remedial action on the most important life stages. The difficulty of tracking small animals over vast distances is immense — attached telemetry devices that measure and transmit data are currently the only feasible method of following adult eels across the ocean. The miniaturization of transmitters in the coming years should advance knowledge considerably. More information on tropical eel species is also needed, as we know even less about them — a new species was even discovered recently in the Philippines (*Anguilla luzonensis*) — and different factors will affect their survival.

*Eels* is a solid introduction to global *Anguilla* species. It provides a convincing argument that eels should be preserved because of their unique life cycle, and their economic and cultural importance. To restore and manage eel populations worldwide, we need a deeper understanding of their life history. ■

**Kim Aarestrup** is a senior scientist in the National Institute of Aquatic Resources at the Technical University of Denmark, 8600 Silkeborg, Denmark.  
e-mail: kaa@aqu.dtu.dk

As biologist Ken Thompson explains in *Do We Need Pandas?*, conserving rare species does not really benefit people. If you care about nature because of its usefulness to humanity, pandas are a luxury item — and so are most other rare species. The money that is spent on saving them could be better applied by protecting ecosystems that provide us with food, timber, clean water, a liveable climate and flood protection.

If one's aim is to prevent extinctions, as in much of traditional conservation, then identifying and fussing over endangered species is the best way forward. If one sees the environment as a source of services, as Thompson does, the more sensible course is to "conserve the fabric of whole ecosystems, and let the rare species look after themselves".

The reason, he explains, is partly because rare species are too sparse to significantly influence the functioning of an ecosystem. They are thus unlikely to be essential for the continued provision of ecosystem services.

Thompson traces conservation scientists' failed attempts to prove that biodiversity is inherently good for ecosystems. First, the results of these experiments — typically using small plots containing manipulated numbers of plants — were not what they were cracked up to be. Yes, more-diverse ecosystems were more productive on average. But this was not a result of their variety alone — it was because they were also more likely to include the most productive plant, monocultures of which could be even more productive. Second, experimenters defined productivity in terms of turning sunlight into biomass.

Yet growth need not tally with value: in lakes, high productivity often means more algae and fewer fish.

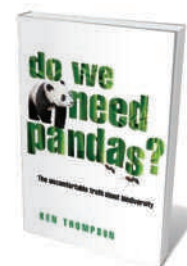
Thompson proposes

that we give up the goal of maximizing biodiversity. Instead, we should focus on saving whole ecosystems that are useful for humanity. In the process of conserving such areas, biodiversity will be protected anyway, as a sort of bonus prize.

But by putting the focus only on what nature can do for us, Thompson leaves open the possibility that ecosystems that do not deliver sufficient services might be thrown out, with all the biodiversity that they contain. He admits that society has benefited from the turning over of forests and wetlands to agriculture: "It is only because of such conversion that you and I have enough to eat." But he does not support conversion of any of the remaining wild habitat. Others disagree: some economists might argue that a particular wild patch would provide better services to humanity as pasture or plantation. This is the peril of the ecosystem-services model. Hitch your wagon to it, and when conversion provides better services than protection, your biodiversity bonus is cancelled.

Despite his book's provocative title, Thompson does not claim that we don't need pandas. Like most ecosystem services enthusiasts, he is keen to have his economic pragmatism and his emotional love of nature too. Letting the panda go extinct would be "a profound failure for our stewardship of the natural world", he feels. But he cannot have it both ways. If the ecosystems in which pandas live do not provide economically valuable services to humanity, then it is goodbye panda. ■

**Emma Marris** writes for Nature from Columbia, Missouri.



**Do We Need Pandas? The Uncomfortable Truth About Biodiversity**  
KEN THOMPSON  
Green Books: 2010.  
160 pp. £9.95

**"We should focus on saving whole ecosystems that are useful for humanity."**

➔ **NATURE.COM**  
For a review on protecting the panda, see:  
[go.nature.com/3dk65d](http://go.nature.com/3dk65d)





Some illegal drugs such as heroin have been demonized only since the mid-nineteenth century.

## NEUROSCIENCE

# Opiates for the people

W. F. Bynum applauds an open-minded exhibition on the history of recreational drugs.

The exhibition *High Society* may not alter your mind, but it will broaden your thinking. The ubiquity of recreational drug use in all times and all cultures is the focus of this show at London's Wellcome Collection. On offer is a varied combination of artefacts, artworks, books and videos tackling pharmacology, the drug trade, self-experimentation, collective intoxication and the ethics of abusing these substances.

Opium and its products pop up again and again: the largest object in the exhibition is a gigantic opium pipe from the late nineteenth century. The dominance of opium is unsurprising, as the poppy was cultivated for its sap at least as early as 3000 BC. Both opium and alcohol have been

**High Society: Mind-Altering Drugs in History and Culture**

Wellcome Collection, London.  
Until 27 February 2011.

drug available for testing *in situ*, at the Wellcome Collection's café.

One man's stimulant is another man's poison. This exhibition takes a non-judgemental stance, making no distinction between legal thrills — such as those delivered by coffee, alcohol and tobacco — and illegal ones. In any case, as the historical and anthropological arcs of the displays show, the legal issues are time- and culture-

intimately connected with humanity for millennia. However, cocaine, mescaline, LSD, tobacco and even betel nuts also get their due. Coffee is here, too — the only

dependent. Opium was demonized in the West only from the mid-nineteenth century; marijuana is always contentious; and many moral or health crusaders would today ban tobacco, alcohol or both.

Running throughout the exhibition is the striking documentation of the fine line separating pleasure and pain, ecstasy and vacant dependency. This is most explicit on a wall of 19 photographs by Tracy Moffat from 1999, entitled *Laudanum*. The images — depicting a woman and her maid experiencing hallucinations in various guises and poses, some erotic, some melancholic and some downright disturbing — convey a sense of what it would be like to witness such an event.

Context is important in experiencing mind-altering drugs. Social cohesion and rites of passage are linked to some hallucinogens, as the exhibition's images and films reveal. Ayahuasca, a psychoactive brew made from *Banisteriopsis* vines, is shown being drunk by the Tucano people of the Colombian Amazon. Peyote, a cactus containing mescaline among a range of alkyls, is dried and chewed by the Huichol and other indigenous peoples of Mexico. The resin from the bark of the virola tree is snorted like snuff during a ceremony in a Venezuela palm festival. These colourful rituals stand in contrast to less convivial Western gatherings, including a counterculture photograph of '4:20 Day', when 10,000 people gathered on 20 April 2008 for a mass 'smoke-in' of cannabis at the University of Colorado, Boulder, to protest against the drug's illegal status.

Medical science also gets a look in. Swiss chemist Albert Hofmann's description of LSD is on display, as well as classic texts about clinical uses of opium and its derivatives. In a recorded interview, British neuroscientist Barry Everitt describes the latest research on neural networks that is helping us to understand addiction and prevent relapse. He explains how addiction may be a learned behaviour that is intimately tied to memory retrieval.

A display on the economics of drugs shows some striking figures. The annual expenditure by the US government on its 'War on Drugs' equals the annual worldwide income of the Roman Catholic Church, or the amount Americans spend each year on complementary medicine. And each of these exceeds the yearly international market for antidepressants.

The centre of gravity of this rewarding exhibition seems to lie in the 'swinging sixties'. If you didn't experience that era first hand, as I did, this is your chance to find out what all the hype was about. ■

W. F. Bynum is emeritus professor of the history of medicine at University College London, UK.  
e-mail: w.bynum@ucl.ac.uk

# CORRESPONDENCE

## Economic growth: indicators not targets

Peter Victor questions the merits of economic growth in developed countries (*Nature* 468, 370–371; 2010). In such discussions, it is important to avoid confusing indicators with optimization targets. An indicator that may be useful for evaluating an economy could be harmful when used as a target to improve the state of the economy.

An economic indicator, such as gross domestic product (GDP) or the genuine progress indicator (GPI), is a number that quantifies a particular aspect of an economy. Indicators are useful for comparing different economies or for monitoring development. But they are overly simplistic in that they ignore all non-quantifiable aspects of living.

This flaw becomes crucial when an indicator is turned into an optimization target. Politicians will quickly identify and exploit mechanisms that are likely to increase the indicator, even if there is no benefit for society.

Measures that would not even be considered in the absence of a specific optimization target can then become political priorities when that target is adopted. Public debt is one such indicator that has recently become a high-priority optimization target in many European countries, despite wide recognition of the socially negative effects of the cuts that are needed to reduce it.

This will happen to any indicator, including the GPI, which takes into account social and environmental factors as well as economic ones. Economists and politicians must accept that no single number can safely be optimized. Several indicators that concentrate on different

aspects of society need to be used in parallel, and any measure that improves one while decreasing another must be recognized as a compromise between conflicting goals.

**Konrad Hinsén** *Centre de Biophysique Moléculaire (CNRS), France.*

*konrad.hinsen@cnrs-orleans.fr*

## Economic growth: enough is enough

There is substantial evidence that further economic growth in wealthy nations is neither sustainable nor desirable. It is indeed time, as Peter Victor writes (*Nature* 468, 370–371; 2010), to answer key questions about what a non-growing economy would look like in practice. We need a new macro-economics for sustainability, and we need it now.

On 17 November a report was released in the United Kingdom, entitled *Enough is Enough: Ideas for a Sustainable Economy in a World of Finite Resources* (see [go.nature.com/hv52np](http://go.nature.com/hv52np)). The report brings together the ideas generated at the first Steady State Economy Conference held in June this year in Leeds, UK. It discusses policy proposals in ten key areas needed to achieve a no-growth economy. Proposals include policies to limit resource use, reduce income inequality, reform the monetary system, change consumer behaviour, restructure business, secure full employment and improve the way in which we measure progress.

A growing number of economists, scientists and policy-makers are beginning to understand the urgent need for an economic model based on stability instead of growth (see [go.nature.com/f8ig8s](http://go.nature.com/f8ig8s)). A combination of further research into the steady-state model and bold action to turn this model

into government policy is required to achieve well-being for everyone within ecological limits.

**Daniel W. O'Neill** *Center for the Advancement of the Steady State Economy, Leeds, UK.*

*dan\_oneill@steadystate.org*

## Coordinate green growth

Green economic growth needs a shared sense of direction if it is to lead to a more sustainable future under climate change. Studies on green innovation and societal transformation show that uncoordinated initiatives are unlikely to be an effective way “to get the ball rolling and to ‘learn by doing’” (*Nature* 468, 477; 2010).

First, socio-technical transformations, such as the transition from fossil fuels to renewable-energy sources, will require several decades to complete. Speeding up this process needs focus and coordination at the international level.

Second, the learning curves for creating energy-efficient and renewable-energy technologies are global. Here, coordination will be necessary to determine cost reductions and to increase performance.

Third, green growth calls for major shifts in the way in which economies are organized. It is not trivial to align the interests of fossil-fuel-intensive incumbent industries and their supporting power structures with the interests of emerging ‘green’ industries. Again, coordination will be necessary to overcome the resistance to change in incumbent production and consumption systems.

**Floortje Alkemade, Marko Hekkert** *Utrecht University, the Netherlands.*  
*f.alkemade@geo.uu.nl*

## A slip in the date of DNA's discovery

In her review of Anna Ziegler's play *Photograph 51*, Josie Glausiusz refers to DNA's “discovery” in 1953 (*Nature* 468, 375; 2010), when this was in fact the year its structure was solved. The molecule itself was discovered almost a century earlier.

It was a young Swiss physician, Friedrich Miescher, who stumbled on DNA in 1869, naming it nuclein. He realized that it chemically defines the nucleus — an enigmatic organelle at that time — and identified the molecule in a wide variety of cell types, including germ cells. He determined DNA's elementary composition and basic biochemical properties, and suggested that it could be important in cell proliferation, realizing it was synthesized before cell division.

Miescher developed theories on the basis of these findings to explain DNA's function in terms of fertilization and heredity, even proposing how macromolecules might encode information. His work also stimulated others to investigate DNA and its function.

Miescher should therefore be remembered not just as the discoverer of DNA, but also as the founder of molecular genetics.

**Ralf Dahm** *Institute of Molecular Biology, Mainz, Germany.*  
*r.dahm@imb-mainz.de*

### CONTRIBUTIONS

Items for Correspondence may be submitted to [correspondence@nature.com](mailto:correspondence@nature.com) after consulting the author guidelines at <http://go.nature.com/cmchno>. They should be no longer than 350 words.

# Allan Sandage

## (1926–2010)

Astronomer who measured expansion rate of the Universe.

Allan Rex Sandage was one of the most prolific and influential astronomers of the second half of the twentieth century. Edwin Hubble and Walter Baade both left their scientific papers to him, and he continued the work of these two giants with spectacular results — including the first good estimates of the Hubble constant and the age of the Universe.

These are his best known achievements, but they comprised only a small part of his publications, which number in excess of 500. In 1956, Sandage became a staff member of the Mount Wilson and Palomar observatories in California; his retirement almost 50 years later did not stop him from working until the very end — his last paper is still in the press. He died, aged 84, on 13 November 2010 at his home in San Gabriel, California, from pancreatic cancer. He is survived by his caring astronomer wife Mary Sandage and his two sons David and John.

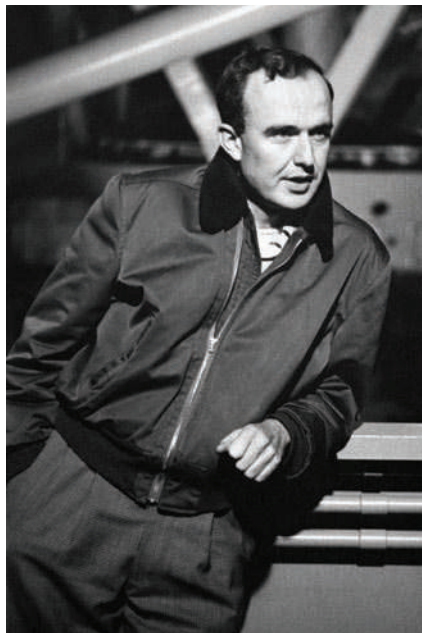
Sandage was born on 18 June 1926 in Iowa City, Iowa, as the only child of a powerful father — a professor of advertising — and loving mother. He received his BA in 1948 from the University of Illinois at Urbana-Champaign and then moved to the California Institute of Technology in Pasadena, California, for his PhD. Here he became Hubble's assistant in 1952, a year before completing his degree and Hubble's sudden death. He chose Baade as his thesis adviser, and Baade went on to teach him all the intricacies of the observing techniques of the time. This made Sandage, who spent roughly 2,000 nights at the telescope during his lifetime, an outstanding observer. He published the famous book *The Hubble Atlas of Galaxies*, followed by a further two colossal atlases containing some of the best ground-based pictures of galaxies ever taken.

### EXPANDING THE UNIVERSE

From the very beginning of his career, Sandage was a star. His 1953 PhD thesis reversed the thinking of the day that faint 'main sequence' stars started their lives as red giants; his measurements of the M3 globular cluster led to the conclusion that the opposite order was correct. This was a revolution in the understanding of stellar evolution. Sandage continued to work on determining the distances and ages of clusters, as well as on the properties of their variable stars — the RR Lyrae and Cepheid stars — all his life. He was one of the first to use supernovae

to measure very large distances, and led a Hubble Space Telescope team to calibrate their luminosity.

By 1958, Sandage had massively revised Hubble's estimates of galactic distances, increasing them by a factor of about 7. From this he determined an expansion rate of the Universe, otherwise known as the Hubble constant, of about 75 kilometres per second per megaparsec, and an age of the Universe of around 13 billion years. Today's best estimates are, remarkably, essentially the same — albeit with smaller errors.



Perhaps his seminal work was his 1961 paper 'The ability of the 200-inch telescope to discriminate between selected world models', which has become the basis of modern observational cosmology. In it, Sandage calculated what the past and future would look like under different models of an expanding Universe and predicted the consequences for an observer.

On the back of these predictions, Sandage single-handedly mounted a giant programme to extend the Hubble diagram, which charts redshifts of galaxies (a measure of how fast they are moving away from Earth) against their relative distances. This became the crucial piece of evidence that helped to dispel doubts about whether very large redshifts are really caused by cosmic expansion, or by some as-yet-unknown physics. Such doubts,

fostered by Hubble himself, became prevailing when the stunningly large redshifts of quasars were discovered in 1963. Sandage's work was instrumental in settling the debate in favour of an expanding Universe.

### QUIET DISCOVERIES

Sandage was actively involved in the discovery of quasars — enigmatic objects first detected by radio astronomers in the late 1950s. Sandage provided some of their first optical identifications and the first spectrum. He also co-discovered radio-quiet quasars, and showed that these objects greatly outnumber their radio-loud counterparts. In 1963, he co-wrote a paper on violent processes in galactic centres that anticipated today's explanation of quasars as being very distant galactic nuclei powered by black holes.

Sandage's most cited paper is from 1962, in which he theorized how the pancake-shaped Milky Way was formed by the collapse of a spherical gas cloud. This work remains the basis of modern theories of galaxy formation.

Sandage also authored several essays on the history of modern astronomy, including a monumental history of the Mount Wilson Observatory.

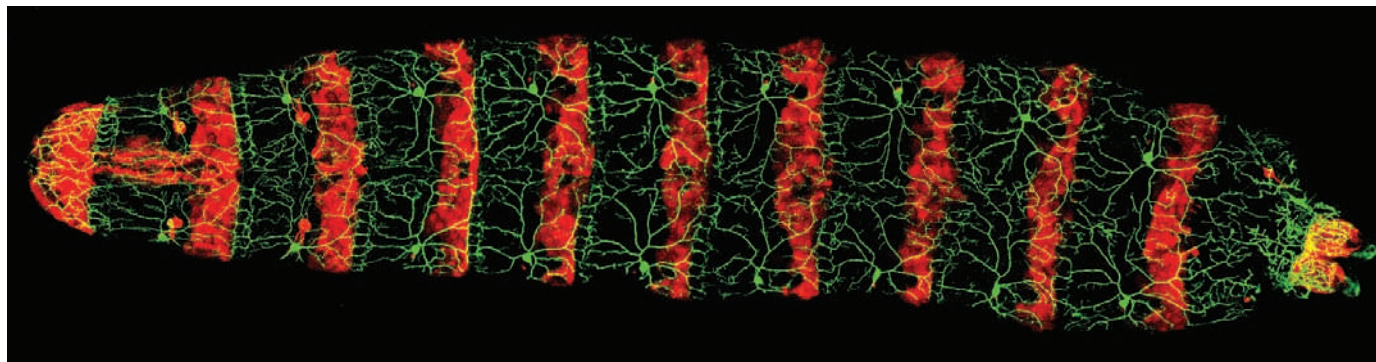
In mid-career, Allan became deeply concerned about the meaning of life. He studied the Bible and spoke in public about science and religion as "two separate closets in the same house". In the end he highly valued Christian philosophy, but did not find faith. He resolved to work to the limit of exhaustion. Some people thought that he was ambitious, but his drive came from his conviction that work was the only meaningful human activity. For Allan, life was not about fun.

Allan sometimes called himself a curmudgeon. But in his social life he was radiant with charm and wit; his after-dinner talks could make his audience explode with laughter. He loved books and was a fan of opera. His favourite time was sitting at the dark telescope, as he did for so many nights of his life, taking long exposures while the dome reverberated with the music of composer Richard Wagner. ■

**Gustav A. Tammann** is professor emeritus at the University of Basel, CH-4003 Basel, Switzerland, and was a co-author of *Allan Sandage* for almost 50 years.  
e-mail: g-a.tammann@unibas.ch

ESTATE OF F. BELLO/SPL





CHUN HAN, UCSF

**Figure 1 | A *Drosophila* larva.** The photoreceptors that mediate light avoidance are labelled green, with the posterior epidermal cells of each segment labelled red.

## NEUROSCIENCE

# Feel the light

**How is light perceived? The answer that might immediately come to mind is, through the eyes. Fly larvae, however, can ‘feel’ light using specialized neurons embedded under the cuticle encasing their bodies. [SEE ARTICLE P.921](#)**

PAUL A. GARRITY

Light perception is a highly useful skill. Like other animals, we humans rely on vision to navigate, to locate food and mates, and to avoid predators. But biological applications of light perception go well beyond vision — from basic light-avoidance to circadian rhythms<sup>1</sup>. What’s more, photoreceptive cells are located not only in the eyes, but also in various non-ocular locations, ranging from the skin in molluscs<sup>2</sup> to the hypothalamus deep within a bird’s brain<sup>3</sup>. Even overtly eyeless animals, such as the soil-dwelling nematode *Caenorhabditis elegans*, possess photosensitive neurons that help them to avoid the daylight<sup>4</sup>.

In this issue, Xiang *et al.*<sup>5</sup> (page 921) extend the analysis of non-ocular photoreception to the fruitfly *Drosophila melanogaster*. They describe a set of dermal photoreceptors that, surprisingly, had previously escaped notice in this well-studied organism, and uncover a molecular mechanism of phototransduction that has not been previously encountered in the fly.

The lives of *Drosophila* larvae are highly focused on burrowing. As they increase in size in preparation for adulthood, the larvae must feed ravenously, and immersing themselves in the nutritious goo of rotting fruit is an excellent way to access a surfeit of calories in a hurry. In addition, the larvae are highly vulnerable to predation when exposed, because their squishy bodies crawl along rather slowly. Burrowing helps them keep out of harm’s way. Tunnelling therefore provides a singular solution for the larva’s need for both feeding and defence.

One of the cues that fly larvae use to orient

the crucial drive towards the interior is light. Young fly larvae are highly photophobic, and this behaviour involves a pair of primitive eye-like structures inside the larva, near its anterior<sup>6</sup>. These structures, called Bolwig organs, resemble the compound eye of the adult fly in many respects, including their expression of light-sensing rhodopsin pigments<sup>7</sup>. But whereas Bolwig organs can lead larvae out of the light, their anterior location raises a potential problem: once the larval anterior is submerged, the light-driven force for burrowing should diminish. This could leave the larva in the awkward position of posterior exposure, like the proverbial ostrich with its head in the sand.

Xiang *et al.*<sup>5</sup> elegantly attack this ethological conundrum. By genetically ablating the Bolwig organs, the authors show that, although Bolwig neurons are crucial for avoiding low light intensities, the requirement for these cells wanes as light intensities approach those of direct sunlight — around 1 mW per mm<sup>2</sup>.

This observation suggests that flies contain additional photoreceptors. Suspecting that these photoreceptors could be analogous to dermal photoreceptors described in other creatures<sup>8,9</sup>, the authors systematically scanned the sensory neurons along the larval body wall for physiological responses to light. They note that one particular set of sensory neurons — the class-IV da neurons — is strongly activated by light (Fig. 1). Satisfyingly, these neurons remain light responsive even when grown in isolation in culture, confirming their intrinsic light sensitivity. Flies therefore contain dermal photoreceptors.

Do these dermal photosensors mediate avoidance of high-intensity light? Xiang and colleagues’ genetic-ablation experiments indicate that they do. Killing class-IV da neurons significantly reduced avoidance at all light intensities. Somewhat surprisingly, however, killing just these neurons, but leaving Bolwig organs intact, dramatically decreased responses to high-intensity light. This suggests that, rather than having overlapping, redundant functions, these two classes of photosensors drive behaviour over different ranges of light intensity.

Intriguingly, previous studies<sup>10–12</sup> have shown that class-IV da neurons also participate in aversive responses to noxious heat and mechanical force. Together with the present results, it seems that these neurons serve as multi-purpose triggers of avoidance. On activation by light, they may provide that extra jolt the larva needs to ensure that its entire body is fully protected.

More surprises were in store when Xiang *et al.* probed the phototransduction machinery of class-IV da neurons. Activation of these cells by light was unaffected when the researchers eliminated proteins on which other fly photoreceptors depend, such as the photon-detecting rhodopsins. Instead, it depended on another G-protein-coupled receptor, Gr28b.

Initially classified as a gustatory receptor, LITE-1 — a nematode relative of Gr28b — was recently discovered<sup>4,13</sup> to mediate phototransduction in *C. elegans*. But although Gr28b and LITE-1 are related, initial evidence suggests differences in the phototransduction pathways in which they are involved. LITE-1

acts through the cyclic nucleotide cGMP to activate cyclic-nucleotide-gated ion channels<sup>4</sup>. Xiang and colleagues' pharmacological data suggest, however, that these channels might not be required for Gr28b activity. Instead, phototransduction in the class-IV da neurons relies on a member of the TRP family of cation channels, TRPA1.

*Drosophila* TRPA1 is known to act as a molecular sensor of temperature<sup>14–16</sup> and of reactive electrophiles<sup>17</sup>, such as the wasabi ingredient allyl isothiocyanate. TRPA1 is also distantly related to the TRP channels that act downstream of rhodopsins in the fly, although this protein was not previously implicated in photodetection. Precisely how TRPA1 cooperates with Gr28b to mediate phototransduction remains to be determined, but activation by G-protein signalling seems a reasonable possibility.

A key issue this paper<sup>5</sup> raises is the mechanism(s) by which proteins such as LITE-1 and Gr28b participate in phototransduction. When misexpressed, LITE-1 can make cells photosensitive<sup>4,13</sup>, suggesting that it could participate in photon detection. Whether Gr28b shares this capability is not known, but it raises the question of how photons might interact with these molecules, and whether the mechanisms used by rhodopsins might have some relevance here. As Gr28b and LITE-1 have additional relatives in flies, worms and other invertebrates, related pathways may be deployed elsewhere in these animals. From a broader evolutionary perspective, one wonders about the origins of these light sensors and the extent to which their functional analogues may occur in other present-day organisms, but have simply escaped our notice — as was the case for so long in *Drosophila*. ■

**Paul A. Garrity** is in the Department of Biology and National Center for Behavioral Genomics, Brandeis University, Waltham, Massachusetts 02454-9110, USA.  
e-mail: pgarrity@brandeis.edu

1. Yau, K. W. & Hardie, R. C. *Cell* **139**, 246–264 (2009).
2. Pankey, S. et al. *J. Comp. Physiol. B* **180**, 1205–1211 (2010).
3. Halford, S. et al. *Curr. Biol.* **19**, 1396–1402 (2009).
4. Liu, J. et al. *Nature Neurosci.* **13**, 715–722 (2010).
5. Xiang, Y. et al. *Nature* **468**, 921–926 (2010).
6. Mazzoni, E. O., Desplan, C. & Blau, J. *Neuron* **45**, 293–300 (2005).
7. Sprecher, S. G. & Desplan, C. *Nature* **454**, 533–537 (2008).
8. Steven, D. M. *Biol. Rev. Camb. Phil. Soc.* **38**, 204–240 (1963).
9. Millott, N. *Symp. Zool. Soc. Lond.* **23**, 1–36 (1968).
10. Tracey, W. D. Jr, Wilson, R. I., Laurent, G. & Benzer, S. *Cell* **113**, 261–273 (2003).
11. Hwang, R. Y. et al. *Curr. Biol.* **17**, 2105–2116 (2007).
12. Zhong, L., Hwang, R. Y. & Tracey, W. D. *Curr. Biol.* **20**, 429–434 (2010).
13. Edwards, S. L. et al. *PLoS Biol.* **6**, e198 (2008).
14. Rosenzweig, M. et al. *Genes Dev.* **19**, 419–424 (2005).
15. Hamada, F. N. et al. *Nature* **454**, 217–220 (2008).
16. Viswanath, V. et al. *Nature* **423**, 822–823 (2003).
17. Kang, K. et al. *Nature* **464**, 597–600 (2010).

## ASTROPHYSICS

# A strange ménage à trois

**The two Magellanic Clouds may have joined our Milky Way quite recently. It turns out that this trio of galaxies is remarkably unlike most other galaxy systems — both in the luminosity of the clouds and in their proximity to the Milky Way.**

SIDNEY VAN DEN BERGH

**W**e are all Copernicans now. So we expect to be living in a typical galaxy in a normal neighbourhood. The first of these expectations is fulfilled: our Milky Way is a relatively normal giant galaxy with fairly loosely wound spiral arms (Hubble type Sbc), or perhaps a spiral giant with a central bar-shaped region of stars (SBbc). But the second expectation is not fulfilled: the Galactic neighbourhood is unusual and quite different from what might have been expected. True, the Local Group that we belong to is a small cluster, like many others in nearby regions of the Universe. However, the nearest neighbours to our home Galaxy have been observed to exhibit remarkable peculiarities. Two papers, one in *Monthly Notices of the Royal Astronomical Society*<sup>1</sup> and the other a recent preprint<sup>2</sup>, now reinforce these observations.

For most galaxies, including Andromeda<sup>3</sup>, the nearest neighbours are elliptical galaxies or lenticulars (an intermediate type between an elliptical and a spiral galaxy), whereas the more distant companions are spirals with loosely

bound spiral arms or galaxies with an irregular shape. However, the Milky Way's two closest big companions, the Large Magellanic Cloud (LMC; Fig. 1) and the Small Magellanic Cloud (SMC), are irregular galaxies. This anomaly suggests<sup>4</sup> that the Magellanic Clouds might not always have been close satellites of the Galaxy, but instead that they might be objects formed in the outer reaches of the Local Group and that just happen to be passing close to the Milky Way at present. Recent calculations<sup>5</sup> suggest that there is a probability of about 72% that the Magellanic Clouds were accreted onto the Milky Way within the past billion years, and a roughly 50% probability that they were accreted together.

The second anomaly among the closest large companions to our Galaxy is that the LMC is extraordinarily luminous for a Magellanic-like irregular galaxy. In nearby regions of the Universe, there are only two Magellanic-like irregular galaxies (NGC 4214 and NGC 4449) that even come close to rivalling the LMC in luminosity. In other words, the LMC seems to be close to the upper luminosity limit for Magellanic-like irregular galaxies. This is



**Figure 1 | The Large Magellanic Cloud.** Calculations by James and Ivory<sup>1</sup> and by Liu et al.<sup>2</sup> suggest that the a priori probability of the Milky Way having a nearby satellite galaxy as luminous as the Large Magellanic Cloud is very low.

L. DODD/SPL



important, because there is a fundamental morphological difference between spirals and Magellanic-like irregular galaxies: spirals, which have a large range of luminosities, all have nuclei, whereas Magellanic irregulars, which are mainly quite faint, do not. It should be emphasized that this upper luminosity limit applies only to Magellanic irregulars and not to the peculiar, chaotic irregular galaxies that might have been formed during the collisions or mergers of massive ancestral galaxies.

In 1969, Erik Holmberg<sup>6</sup> searched for the satellites of nearby galaxies on the photographic prints of the Palomar Sky Survey. Surprisingly, he found that bright satellite galaxies like the Magellanic Clouds are quite rare. This conclusion is now strengthened and confirmed by the work of James and Ivory<sup>1</sup> and that of Liu and colleagues<sup>2</sup>. James and Ivory used narrow-spectral-band imaging of 143 luminous spiral galaxies comparable to the Milky Way to search for star-forming companions. They concluded that luminous, star-forming satellite galaxies resembling the Magellanic Clouds are quite uncommon, and that our home Galaxy is unusual, both for the luminosity and the proximity of its two brightest satellites (the Magellanic Clouds).

A different approach was employed by Liu *et al.*<sup>2</sup>, who used the enormous database provided by the Sloan Digital Sky Survey to search for satellite galaxies, around Milky-Way-like host galaxies, that have luminosities similar to those of the Magellanic Clouds and that are located within a distance of 150 kiloparsecs of their apparent host galaxy; the LMC and the SMC are only 50 and 60 kiloparsecs, respectively, away from the Milky Way. For 22,581 Milky-Way-like hosts, Liu *et al.* found that 81% have no satellites as bright as the Magellanic Clouds, 11% have one such satellite, and only 3.5% host two such galaxies. As Edwin Hubble<sup>7</sup> said many years ago, "The fact that the [G]alactic system is a member of a group is a very fortunate accident." That the Galaxy should have an irregular companion as luminous as the Large Magellanic Cloud is almost a miracle. ■

**Sidney van den Bergh** is at the Dominion Astrophysical Observatory, Herzberg Institute of Astrophysics, National Research Council of Canada, Victoria, British Columbia V9E 2E7, Canada.  
e-mail: [sidney.vandenbergh@nrc-cnrc.gc.ca](mailto:sidney.vandenbergh@nrc-cnrc.gc.ca)

- James, P. A. & Ivory, C. F. *Mon. Not. R. Astron. Soc.* (in the press); preprint at <http://arxiv.org/abs/1009.2875> (2010).
- Liu, L. *et al.* Preprint at <http://arxiv.org/abs/1011.2255v2> (2010).
- Einasto, J. *et al.* *Nature* **252**, 111–113 (1974).
- van den Bergh, S. *Astron. J.* **132**, 1571–1574 (2006).
- Busha, M. T. *et al.* Preprint at <http://arxiv.org/abs/1011.2203v2> (2010).
- Holmberg, E. *Ark. Astron.* **5**, 305–343 (1969).
- Hubble, E. *The Realm of the Nebulae* (Yale Univ. Press, 1936).

## DRUG DISCOVERY

# How melanomas bypass new therapy

The promise of an exciting new drug that inhibits the mutant B-RAF protein in skin cancer is marred by the fact that most patients relapse within a year. Fresh data hint at how such resistance emerges. **SEE LETTERS P.968 & P.973**

DAVID SOLIT & CHARLES L. SAWYERS

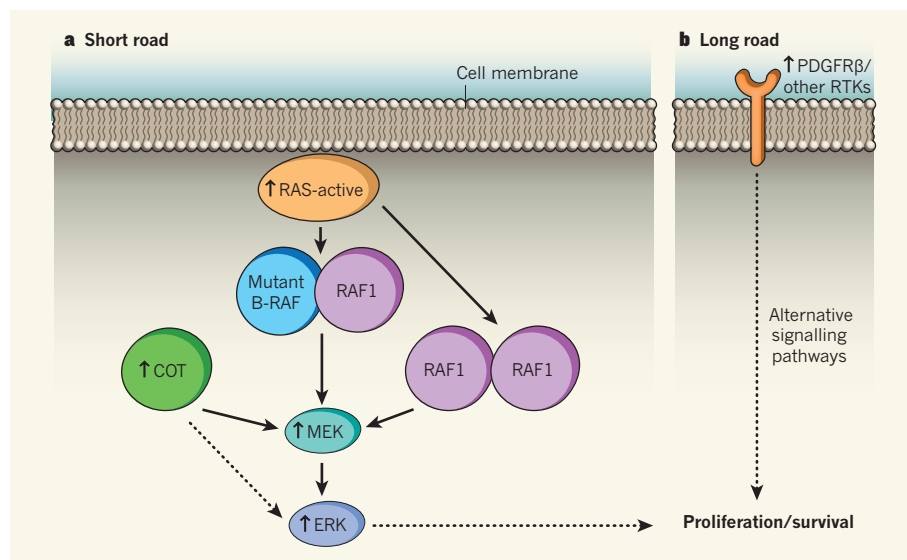
Activating mutations in the kinase enzyme B-RAF, a regulator of cell proliferation and survival, occur in 60% of patients with melanoma, an often fatal form of skin cancer<sup>1</sup>. This discovery — made through large-scale sequencing of cancer genomes — provided the rationale for the development of PLX4032, a drug that inhibits B-RAF and has shown remarkable clinical activity in patients with B-RAF-mutant melanomas<sup>2,3</sup>. PLX4032 is under study in a phase III clinical trial, which could result in its approval within a year. But this clinical success is only transient: resistance to PLX4032 develops quickly, typically within 8–12 months following treatment, even in patients whose tumours seem to have disappeared on radiographic scans. Two papers<sup>4,5</sup> in this issue provide the first clues as to why.

Resistance to cancer drugs — like that to antibiotics — is an unfortunate, yet familiar, story. The anticancer drugs erlotinib, crizotinib

and imatinib, which also function by inhibiting kinases, are effective treatments for lung cancer, leukaemia and gastrointestinal stromal tumours. Nonetheless, the long-term efficacy of all these compounds is also limited by drug resistance.

Understanding the resistance mechanism can provide clues both for developing improved versions of a drug and for guiding the selection of appropriate drug combinations for therapy. For instance, drug-resistant tumour cells often contain secondary mutations in the target kinase that prevent the drug from binding, while retaining the kinase's full oncogenic activity. The discovery of this mechanism of resistance in chronic myeloid leukaemia accelerated the development of two next-generation inhibitors (dasatinib and nilotinib), which are likely soon to become front-line therapies<sup>6,7</sup>.

Given this precedent, one would expect secondary mutations in B-RAF to be the primary cause of resistance to PLX4032. Shockingly,



**Figure 1 | The short and long roads to PLX4032 resistance<sup>4,5</sup>.** **a**, In cells expressing mutant B-RAF, overexpression of RAF1, or activation of RAS due to N-RAS mutation, results in the formation of B-RAF-RAF1 heterodimers and/or RAF1-RAF1 homodimers, causing resistance to PLX4032. Alternatively, overexpression of COT results in RAF-independent activation of MEK and ERK and thus resistance to PLX4032. In such cells, therefore, PLX4032 resistance is mediated by reactivation of the MAPK/ERK signalling pathway. **b**, Another possibility is that activation of upstream receptor tyrosine kinases (RTKs) such as PDGFR $\beta$  makes MEK activity redundant by triggering downstream effectors of cell transformation through parallel signalling pathways.

however, Nazarian *et al.*<sup>5</sup> (page 973) show that this is not the case. The authors used next-generation sequencing technology to exhaustively examine the genomes of tumour samples from 12 patients with acquired resistance to PLX4032, but found no secondary B-RAF mutations. This observation is even more surprising because, in experimental models<sup>8</sup>, mutations engineered into the 'gate-keeper' threonine residue in the ATP-binding pocket of mutant B-RAF are known to confer PLX4032 resistance.

But aficionados appreciate that oncogenic versions (alleles) of the *B-RAF* gene already violate other sacrosanct kinase rules. Like most other kinases, normal B-RAF forms B-RAF–B-RAF homodimers or heterodimers (often with the related protein RAF1) in response to upstream signals such as activation of the signalling molecule RAS. Oncogenic B-RAF, by contrast, signals as a monomer, activating downstream signals in the absence of upstream input.

PLX4032 selectively shuts down the activity of the mutant B-RAF monomers, potently blocking the growth of tumours that have B-RAF mutations. Paradoxically, this drug also activates downstream signalling in cells lacking B-RAF mutations — both non-cancerous cells and tumour cells with normal B-RAF — through transactivation of the non-drug-bound partner in B-RAF–RAF1 heterodimers and RAF1–RAF1 homodimers<sup>9–11</sup>. This explains both the exquisite specificity of the drug for B-RAF-mutant tumours and, most probably, the complication of low-grade squamous cell carcinomas that can arise from normal skin cells in some patients being treated with PLX4032.

The new papers<sup>4,5</sup> suggest at least three different mechanisms of resistance, all of which share the common theme of 'oncogene bypass'. This concept was first invoked by the discovery<sup>12</sup> that amplification of the tyrosine kinase *MET* bypasses the therapeutic effect of erlotinib on the oncogenic protein EGFR in lung cancer, thereby causing erlotinib resistance.

Following the EGFR/*MET* precedent, Johannessen and colleagues<sup>4</sup> (page 968) methodically examined kinase-coding genes to see if any conferred resistance to PLX4032. They obtained two compelling hits: RAF1, which was expected from earlier work<sup>13</sup> *in vitro*; and COT/TPL2, which, like B-RAF and RAF1, functions upstream of the kinase MEK (B-RAF functions through the MAPK/ERK signalling pathway). In both cases, resistance seems to be due to restored MEK activation in the face of PLX4032 treatment — that is, B-RAF bypass (Fig. 1a). The clinical importance of these hits remains to be defined, because the authors isolated both genes under artificial screening conditions. COT, however, is a particularly intriguing candidate, because it is amplified in a few cell lines with B-RAF mutations that show intrinsic PLX4032 resistance<sup>4</sup>.

Nazarian and co-workers<sup>5</sup> took a different approach. They focused entirely on changes in signalling pathways in 12 matched pairs of drug-sensitive and drug-resistant tumour samples obtained from patients with B-RAF-mutant melanoma and who were treated with PLX4032. In one patient, two independent, drug-resistant subclones of the original tumour acquired new activating mutations in N-RAS, while retaining the original mutant B-RAF allele. As is the case for overexpression of COT and RAF1, expression of the mutant N-RAS allele restored activation of MEK, conferring PLX4032 resistance (B-RAF bypass).

These data suggest that any perturbation that restores MEK activation (RAS mutation, RAF1 or COT overexpression and perhaps other perturbations) has the potential to cause resistance to PLX4032. However, the data obtained from other patients Nazarian *et al.* examined suggest a different mechanism.

In samples from five patients in relapse, the researchers<sup>5</sup> documented increased activation of the receptor tyrosine kinase PDGFR $\beta$ , compared with the baseline. Like increased expression of COT and mutant N-RAS, overexpression of PDGFR $\beta$  was sufficient to confer resistance to PLX4032, albeit independently of MEK activation (Fig. 1b). In the remaining six of the twelve patients the authors studied, the resistance mechanism remains unexplained.

Will the present findings<sup>4,5</sup> have a clinical impact? It is too early to tell, given the limited number of patients examined and the heterogeneity of potential resistance mechanisms. But the fact that restored MEK activation is sufficient to confer resistance in models of B-RAF-mutant melanoma begs the

question of whether, at least in some patients, resistance to PLX4032 can be overcome using MEK inhibitors.

Indeed, if MEK activation is commonly found at relapse, regardless of the mechanism, one could make a strong argument for combined therapy using PLX4032 and a MEK inhibitor to either prevent or delay resistance. The added bonus of this combination is prevention of MEK activation in normal tissues — a side effect of PLX4032 — which could overcome the complication of low-grade squamous-cell carcinomas emerging in some patients. ■

**David Solit and Charles L. Sawyers** are in the Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, New York 10065, USA. e-mail: sawyersc@mskcc.org

1. Davies, H. *et al.* *Nature* **417**, 949–954 (2002).
2. Bollag, G. *et al.* *Nature* **467**, 596–599 (2010).
3. Flaherty, K. T. *et al.* *N. Engl. J. Med.* **363**, 809–819 (2010).
4. Johannessen, C. M. *et al.* *Nature* **468**, 968–972 (2010).
5. Nazarian, R. *et al.* *Nature* **468**, 973–977 (2010).
6. Kantarjian, H. *et al.* *N. Engl. J. Med.* **362**, 2260–2270 (2010).
7. Saglio, G. *et al.* *N. Engl. J. Med.* **362**, 2251–2259 (2010).
8. Whittaker, S. *et al.* *Sci. Transl. Med.* **2**, 35ra41 (2010).
9. Poulikakos, P. I., Zhang, C., Bollag, G., Shokat, K. M. & Rosen, N. *Nature* **464**, 427–430 (2010).
10. Hatzivassiliou, G. *et al.* *Nature* **464**, 431–435 (2010).
11. Heidorn, S. J. *et al.* *Cell* **140**, 209–221 (2010).
12. Engelman, J. A. *et al.* *Science* **316**, 1039–1043 (2007).
13. Montagut, C. *et al.* *Cancer Res.* **68**, 4853–4861 (2008).

#### SOLAR SYSTEM

## Recipe for making Saturn's rings

**Simulations show that the still-mysterious origin of Saturn's vast, icy rings could be explained by the 'peeling' by Saturn's tides of the icy mantle of a large satellite migrating towards the planet. [SEE LETTER P.943](#)**

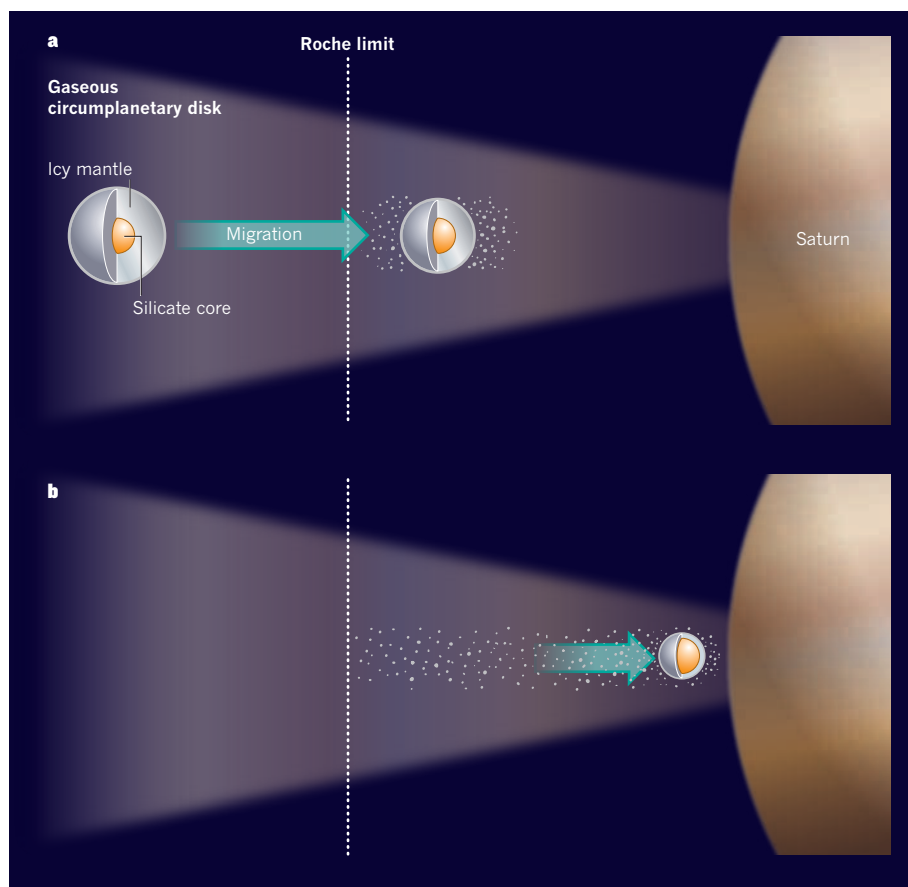
AURÉLIEN CRIDA & SÉBASTIEN CHARNOZ

Since Christiaan Huygens realized in 1655 that the planet Saturn is encircled by a ring, this "jewel of the Solar System" has defied researchers' best efforts to explain its origin. Saturn's rings are made of centimetre-to metre-sized boulders of almost pure water ice<sup>1</sup> — a unique characteristic among Solar System bodies (comets and planetary satellites contain about 50% silicates and metals). The total mass of Saturn's rings is thought to

be equivalent to that of a satellite about 500 kilometres across<sup>2,3</sup>. But how and when the rings formed, and why they are so clean of silicates, is not understood. Several mechanisms for their formation have been proposed<sup>4–6</sup>, but none has provided a convincing explanation for their observed peculiarities. On page 943 of this issue, Canup<sup>7</sup> offers an attractive solution to the problem that answers several questions at once\*.

\*This article and the paper under discussion<sup>7</sup> were published online on 12 December 2010.





**Figure 1 | Canup's model for the formation of Saturn's icy rings<sup>7</sup>.** **a**, A differentiated satellite in the gaseous circumplanetary disk around Saturn migrates towards the planet. When the satellite crosses the planet's Roche limit, its icy mantle starts to be pulled into pieces by the planet's tidal forces. **b**, The silicate core carries on migrating towards Saturn and eventually falls into it, leaving behind the ice boulders that give birth to the rings and icy satellites of the planet.

Saturn's rings are in a location that is dominated by tides. Generally, boulders in circular orbits, such as those making up the rings of Saturn, merge and grow to form larger bodies because of their own gravity — this process is thought to be the way in which planets and asteroids form around the Sun and satellites form around giant planets. But because Saturn's rings are so close to the planet (below the planet's Roche limit, which is 2.5 times the planetary radius), the tidal forces that the planet exerts on them prevent the boulders from accreting material. Just as the Moon's tides stretch the oceans on Earth, Saturn's tides stretch any boulder aggregates in the rings, and separate their constituents. Similarly, if a pre-existing icy body were to be placed within Saturn's Roche limit, it would be destroyed by the planet's tides.

These considerations suggest a simple recipe for making Saturn's rings: take a large body and put it into a close orbit around the planet; tides should then destroy it, and the resulting fragments should form the rings. But what sort of body? A large, differentiated satellite — with a core composed of silicates and iron, and a lighter mantle made of water ice — would be a good choice. The rings could then

be formed simply by 'peeling off' the satellite's icy mantle from the core using the planet's tides as a knife<sup>4–6</sup>. Although appealing, however, this recipe has never been investigated owing to computational limitations.

Enter Canup<sup>7</sup>, who describes the details of a numerical and analytical model for tidal splitting of a differentiated satellite around Saturn. She demonstrates that the planet's tidal forces can indeed be strong enough to take water ice away from the satellite, but not to tear its dense silicate core. But how can the differentiated satellite be brought so close to the planet, within the Roche limit? When could this have occurred? And what mechanism can get rid of the silicate core that has not been broken by the tides? Canup solves these problems using a single phenomenon: planetary migration.

The planets of our Solar System formed 4.5 billion years ago in a disk of gas and dust that surrounded the young Sun for a few million years. At the same time, the satellites of the giant planets formed in gaseous and dusty disks around their hosts. But when a body orbits inside a gaseous disk, its orbit shrinks and the body spirals gradually towards the planet as a result of the gravitational interaction of the body with the surrounding gas. In the

circum-Saturn disk, satellites grew and fell into the planet in this way, until the last generation formed and escaped inward migration because the disk had vanished<sup>8</sup>. Canup suggests that the last migrating differentiated satellite, about the size of Titan (Saturn's largest moon), had its icy mantle pulled to pieces by Saturn's tides as it crossed the Roche limit. Beyond this limit, the satellite's silicate core carried on migrating inwards and eventually disappeared into Saturn, leaving behind the ice boulders that make up the rings (Fig. 1). This process would have produced icy rings about 1,000 times more massive than the rings are today.

This elegant model<sup>7</sup> could provide the missing links between a suite of observational and theoretical results that have changed our understanding of Saturn's rings. Such massive rings would be less sensitive to the darkening effect of meteoroid bombardment<sup>3,9</sup>, thus explaining their brightness today. In addition, because of their mass they should spread more rapidly than the present ones, leading, over the age of the Solar System, to lighter rings like those seen today<sup>7,10</sup>. During this spreading, the material expanding beyond the Roche limit may have given birth to satellites; such a mechanism has recently been proposed for the formation of Saturn's small moons<sup>11</sup>. Indeed, observations made with the Cassini spacecraft have shown that accretion processes are still active at the outer edge of Saturn's rings<sup>12,13</sup> and on the satellites Pan and Atlas<sup>14</sup>. Canup suggests that the inner satellites of Saturn, up to and including Tethys, could also have formed by accretion of spreading ring material beyond the Roche limit.

Canup's model<sup>7</sup> offers, for the first time, a convincing starting point for a consistent theory of the origin of Saturn's rings and satellites. It shows that the rings and satellites are intimately linked, and that Saturn's system, despite being made of ice, is not frozen but is constantly evolving. The origin of the rings and satellites must be understood in the wider framework of models of planet formation, and this work is one step in that direction. One may question whether the specific conditions required for such rings to form have also been met around other Solar System planets and exoplanets. The details and the consequences of the formation of the satellites at the outer edge of the rings, and their outward migration to their present positions, are still to be explored. Establishing these details could change our understanding of Saturn's satellites, and more generally of giant planets and their environments. ■

**Aurélien Crida** is at the *Laboratoire Cassiopée, Université de Nice Sophia-antipolis/CNRS/Observatoire de la Côte d'Azur — BP4229, 06304 Nice Cedex 4, France*. **Sébastien Charnoz** is at the *Laboratoire AIM, Université Paris Diderot/CEA IRFU/CNRS CEA Saclay — SAp Centre de l'Orme Les Merisiers,*

91191 Gif-sur-Yvette Cedex, France.  
e-mail: aurelien.crida@oca.eu

- Nicholson, P. D. *et al.* *Icarus* **193**, 182–212 (2008).
- Esposito, L. W., O'Callaghan, M. & West R. A. *Icarus* **56**, 439–452 (1983).
- Robbins, S. J., Stewart, G. R., Lewis, M. C., Colwell, J. E. & Sremčević, M. *Icarus* **206**, 431–445 (2010).

- Harris, A. in *Planetary Rings* (eds Greenberg, R. & Brahic, A.) 641–659 (Univ. Arizona Press, 1984).
- Dones, L. *Icarus* **92**, 194–203 (1991).
- Charnoz, S., Morbidelli, A., Dones, L. & Salmon, J. *Icarus* **199**, 413–428 (2009).
- Canup, R. M. *Nature* **468**, 943–946 (2010).
- Canup, R. M. & Ward, W. R. *Astron. J.* **124**, 3404–3423 (2002).
- Cuzzi, J. N. & Estrada, P. R. *Icarus* **132**, 1–35 (1998).

- Salmon, J., Charnoz, S. & Crida, A. *Icarus* **209**, 771–785 (2010).
- Charnoz, S., Salmon, J. & Crida, A. *Nature* **465**, 752–754 (2010).
- Beurle, K. *et al.* *Astrophys. J.* **718**, L176–L180 (2010).
- Esposito, L. W., Meinke, B. K., Colwell, J. E., Nicholson, P. D. & Hedman, M. W. *Icarus* **194**, 278–289 (2008).
- Charnoz, S., Brahic, A., Thomas, P. C. & Porco, C. C. *Science* **318**, 1622–1624 (2007).

## CLIMATE CHANGE

# The prospects for polar bears

**Is the polar bear doomed to extinction? Maybe not, according to models of the future extent of Arctic sea ice if greenhouse-gas emissions are curbed. The outlook depends on the ability of policy-makers to act. [SEE LETTER P.955](#)**

ANDREW E. DEROCHE

The projected loss of polar bear sea-ice habitat as a result of a warming climate will dramatically reduce the spatial and temporal extent of that habitat by the end of the twenty-first century<sup>1</sup>. Accordingly, demographic analyses and population-projection models have predicted drastic declines in the polar bear population of the southern Beaufort Sea<sup>2</sup>. Combined with dire predictions from other regions, these findings led to polar bears being listed as 'Threatened' throughout their range under the US Endangered Species Act. However, the listing process took no account of the potential effects of measures to reduce greenhouse-gas emissions, and so reduce anthropogenic warming. Amstrup *et al.*<sup>3</sup> (page 955) provide such analyses and their results are cause for some optimism — if timely action is taken.

Amstrup and colleagues examined projection models of sea-ice loss based on different greenhouse-gas emission scenarios and found that mitigation could greatly improve the conservation status of polar bears into the next century. Reduced emissions, the lower the better, would yield greater abundance and wider distribution of polar bears than the 'business as usual' emission scenario. Lower levels of warming and sea-ice loss would improve the conservation outlook not just for polar bears, but for other Arctic marine species as well. Nonetheless, most emission scenarios resulted in substantial loss of optimal habitats and a major increase in the ice-free period over the more biologically productive

continental shelves. Amstrup *et al.* also reaffirm earlier findings that emission scenarios with high end-of-century levels of carbon dioxide, a major greenhouse gas, will result in great loss of polar bears.

The best-studied polar bear population is in



**Figure 1 | Distribution of the polar bear (*Ursus maritimus*).** In this overlay of the Arctic, darker areas denote higher population density, lighter areas lower density. Amstrup and colleagues' model projections<sup>3</sup> show that implementation of measures to slow global warming improve the prospects for polar bears. But even with such measures, various factors could still conspire to make their future gloomy.

western Hudson Bay, Canada, and it has shown a slow decline due to the lower survival and reproduction that is correlated with reduced sea-ice duration<sup>4</sup>. The possible occurrence of 'tipping points' (where rising temperatures trigger a feedback loop that further drives ice loss<sup>5</sup>), and the summer ice minima<sup>6</sup> of 2007, sparked concerns of sudden and irreversible loss of polar bear habitat. Catastrophic shifts have been noted in many different ecosystems<sup>7</sup>, and in ice-covered Arctic marine ecosystems continued melting of the sea ice could produce episodic changes in predator-prey dynamics and rapid restructuring of the food web<sup>8</sup>.

Amstrup and colleagues<sup>3</sup>, however, found no evidence of a critical temperature that resulted in a tipping point. Rather, they found a linear relationship between global mean surface air temperature and sea-ice extent, and that rapid ice loss could partially reverse. But the authors caution that tipping points could occur in the real world and that the situation could still be dire. Although Arctic marine mammals are well adapted to fluctuating environments, and can tolerate substantial inter-annual change, analyses of polar bear energetics show that rapid ice loss could trigger strong nonlinear declines in survival<sup>9</sup>. Given the low reproductive rates of these animals, even episodic loss of sea ice followed by recovery could have serious effects.

Amstrup *et al.* also note that the best possible outcomes for polar bears include controlling hunting and other factors in an effort to make populations with the expected lower numbers sustainable. But a ban on hunting would be a serious cultural loss for the Arctic's aboriginal people. There are also other caveats. The authors warn that a lack of data across the full polar bear range (Fig. 1) makes predictions of future abundance tenuous. Further, they stress that the relationships between demographics and sea ice are only partly understood, and that geographical differences in population response to even linear declines in sea ice may vary.

Increased risk of extinction is often associated with specialization<sup>10</sup>. So it is not surprising that polar bears — as large predators that rely on sea ice as the platform on which to hunt their seal prey, to mate and to travel — are



especially vulnerable. This paper<sup>3</sup> provides reason to hope that the previous predictions of declines in polar bear populations can be avoided if concerted efforts are made to reduce greenhouse-gas emissions. The threat posed by climate change to biological diversity has been clear for years, however, and calls for carbon sequestration and reduction of emissions to conserve species<sup>11</sup> have largely gone unheeded. Amstrup and colleagues describe polar bears as sentinels of Arctic marine ecosystems and emphasize the importance of sea-ice habitats in the global climate system: both would benefit from greenhouse-gas reduction.

There are few indications, however, that such policies will be implemented in a timely

manner. Globally, 25% of mammalian species are threatened with extinction — with habitat loss and degradation being the main causes<sup>12</sup>. In this context, the plight of polar bears is sadly typical. Their future remains uncertain, but it is now more clearly in the hands of policy-makers. There is cause for optimism, but that requires optimism about our ability to change. ■

**Andrew E. Derocher** is in the Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9, Canada. e-mail: derocher@ualberta.ca

1. Durner, G. M. *et al. Ecol. Monogr.* **79**, 25–58 (2009).

2. Hunter, C. M. *et al. Ecology* **91**, 2883–2897 (2010).
3. Amstrup, S. C. *et al. Nature* **468**, 955–958 (2010).
4. Regehr, E. V., Lunn, N. J., Amstrup, S. C. & Stirling, I. *J. Wildl. Mgmt* **71**, 2673–2683 (2007).
5. Lindsay, R. W. & Zhang, J. J. *Clim.* **18**, 4879–4894 (2005).
6. Comiso, J. C., Parkinson, C. L., Gersten, R. & Stock, L. *Geophys. Res. Lett.* **35**, L01703 (2008).
7. Scheffer, M., Carpenter, S., Foley, J. A., Folke, C. & Walker, B. *Nature* **413**, 591–596 (2001).
8. Higdon, J. W. & Ferguson, S. H. *Ecol. Appl.* **19**, 1365–1375 (2009).
9. Molnár, P. K., Derocher, A. E., Thiemann, G. W. & Lewis, M. A. *Biol. Conserv.* **143**, 1612–1622 (2010).
10. Davies, K. F., Margules, C. R. & Lawrence, J. F. *Ecology* **85**, 265–271 (2004).
11. Thomas, C. D. *et al. Nature* **427**, 145–148 (2004).
12. Schipper, J. *et al. Science* **322**, 225–230 (2008).

## MATERIALS SCIENCE

# Pleated crystals

**A neat study that involves placing colloidal particles on curved oil–glycerol interfaces reveals a new form of crystal defect. The defect is called a pleat, by analogy to the age-old type of fabric fold. SEE LETTER P.947**

**FRANCESCO STELLACCI  
& ANDREAS MORTENSEN**

In 1947, William Bragg and John Nye had a simple yet brilliant idea<sup>1</sup>: to model atoms using bubbles. Bubbles are neatly spherical yet soft, and if many small bubbles, all of the same size, are blown onto the surface of a soapy film, they assemble into a hexagonal array (raft) that mimics the atomic arrangements of flat, two-dimensional crystals. On page 947 of this issue, Irvine and colleagues<sup>2</sup> present an extension of the ‘bubble-raft technique’ to visualize the behaviour of two-dimensional crystals along curved surfaces.

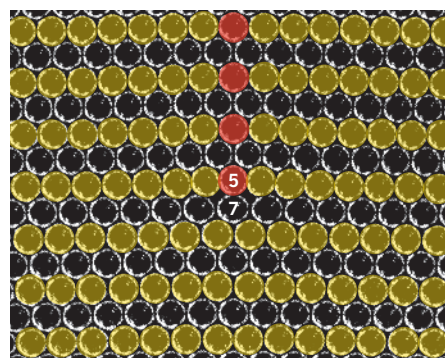
Bubble rafts are ideal tools for studying crystals. They are imperfect: exactly like crystals, they can contain defects such as vacancies, impurities, dislocations and grain boundaries. In two-dimensional crystals, dislocations are point defects (line defects in three dimensions) formed by the termination of a row or column (a plane in three dimensions) of periodically aligned atoms (Fig. 1). Dislocations are important because, if they move, they shift matter by one atomic-lattice spacing about their trajectory; in three dimensions, dislocations are the main mechanism by which metals deform permanently. They are also important in crystalline materials because arrays of dislocations form the building blocks of subgrain boundaries — grain boundaries separating grains that are misoriented by only a few degrees<sup>1</sup>.

The question of whether dislocations can drive the irreversible deformation (plasticity) of materials had been controversial, but was vividly addressed by Bragg and Nye’s original

bubble rafts: when stress was applied to the rafts, dislocations were seen to move and deform them. Movies of bubble rafts offer one of the most eloquent visualizations of dislocations, and are still used today to tackle subtle problems in the plastic deformation of both crystalline and amorphous materials<sup>3</sup>. They can be seen on YouTube, and are often used to illustrate the nature and significance of defects in crystals to students.

Meanwhile, the bubble-raft technique has been extended to three dimensions, and has been widely used to address problems in materials science ranging from melting and other phase changes in crystalline materials<sup>4,5</sup> to the nature of the glassy state<sup>6</sup>. In these studies<sup>4–6</sup>, the three-dimensional bubble rafts are not arrangements of actual bubbles, but instead consist of colloidal crystals, and specialized microscopic techniques are used to study their behaviour. Yet, to date, the bubble-raft principle has seldom been used to study the interplay between the periodicity of a two-dimensional crystal’s lattice structure and the curvature of the surfaces on which it is laid.

In two-dimensional bubble rafts, gravity keeps the bubbles (or crystals) within the flat, liquid-solution surface over which the bubbles are laid. If the bubble raft were transported into space, with gravity now absent, several things would happen. One is that dislocations in the bubble raft might buckle. In one of Bragg and Nye’s original rafts, an extra line of bubbles terminating at a dislocation is visible, and causes the surrounding bubbles to be squeezed against one another (Fig. 1). This squeezing results in an increase in the internal elastic energy stored



**Figure 1 | The bubble raft.** This image shows one of Bragg and Nye’s original bubble rafts<sup>1</sup>, assemblies of bubbles on the surface of a solution consisting of water, glycerine, oleic acid and triethanolamine that can be used to model two-dimensional crystals. Every other horizontal bubble row has been coloured green to facilitate the identification of the column of bubbles that terminates at a dislocation (red). The bubbles that form the bottom of this column (and hence the dislocation) are two opposite disclinations, as predicted by theory. They have a different number of nearest neighbours from bubbles elsewhere: 5 and 7 rather than 6. Horizontal bubble rows tilt downwards slightly to the right of the dislocation. An array of similar dislocations, regularly spaced one over the other, creates a boundary across which the crystal is tilted by a constant angle: this is a two-dimensional subgrain boundary. Irvine and colleagues<sup>2</sup> study the interaction between the curvature of a two-dimensional crystal and the disclinations or dislocations it contains, and find short subgrain boundaries, which they call pleats. (Image reproduced from ref. 1.)

in the raft, which can be relieved if bubbles at the top of the dislocation move out of the plane of the raft<sup>7,8</sup>. Such buckling in turn curves the raft’s surface. In addition, a bubble raft in space will curve more globally, because the liquid-solution surface can seek a shape that minimizes its total surface energy: the surface becomes globally curved. This creates interesting complications for the way in which bubbles pack, because in such two-dimensional arrays there is an interplay between surface curvature and defects.

## NEUROSCIENCE

# Angelman syndrome connections

Neuronal networks in the brain that develop early in life underlie our ability to learn, remember and communicate. Genetic defects that perturb the fine-tuning of such neuronal connectivity can cause disease.

PETER SCHEIFFELE & ASIM A. BEG

Such interplay is well known for disclinations. These are points in the crystal where the rotational lattice symmetry is broken (for dislocations, it is the translational symmetry that is disrupted). At a disclination, an atom has a different number of nearest neighbours from atoms elsewhere (Fig. 1). Disclinations come with much lattice deformation, and hence lead to an excess of elastic energy in the crystal. This energy can be relieved, as for dislocations, by curvature in the surface of the substrate on which the atoms are packed. As a result, disclinations and substrate-surface curvature interact. It is this interaction between curvature and disclinations, or other crystal defects in general, that Irvine and colleagues<sup>2</sup> investigate in their study.

It is not necessary to go into space for gravity's effect to vanish and curve the surfaces: all that is needed is a sufficiently small system. But producing small systems that can be used to study the physics of two-dimensional crystals on curved surfaces is challenging. Irvine and co-workers circumvented this problem by using colloidal particles on curved oil-glycerol interfaces to visualize real, two-dimensional crystals on curved surfaces. They find that the crystals display both five- and seven-neighbour disclinations in curved regions, just as predicted by theory<sup>9</sup>. What's more, they show that two-dimensional subgrain boundaries can form on curved surfaces; they call these pleats, by analogy to the age-old type of skirt. And here, too, they find a correlation with curvature: the pleats' presence, extension and position depend on the curvature. This interplay between pleats and curvature makes the nature and role of pleats interesting, potentially opening up new avenues to explaining some of the many phenomena, both natural and man-made, that are governed by the physics of two-dimensional curved crystals. ■

**Francesco Stellacci and Andreas Mortensen** are in the Department of Materials Science and Engineering, Ecole Polytechnique Fédérale de Lausanne, Lausanne 1015, Switzerland.  
e-mail: francesco.stellacci@epfl.ch

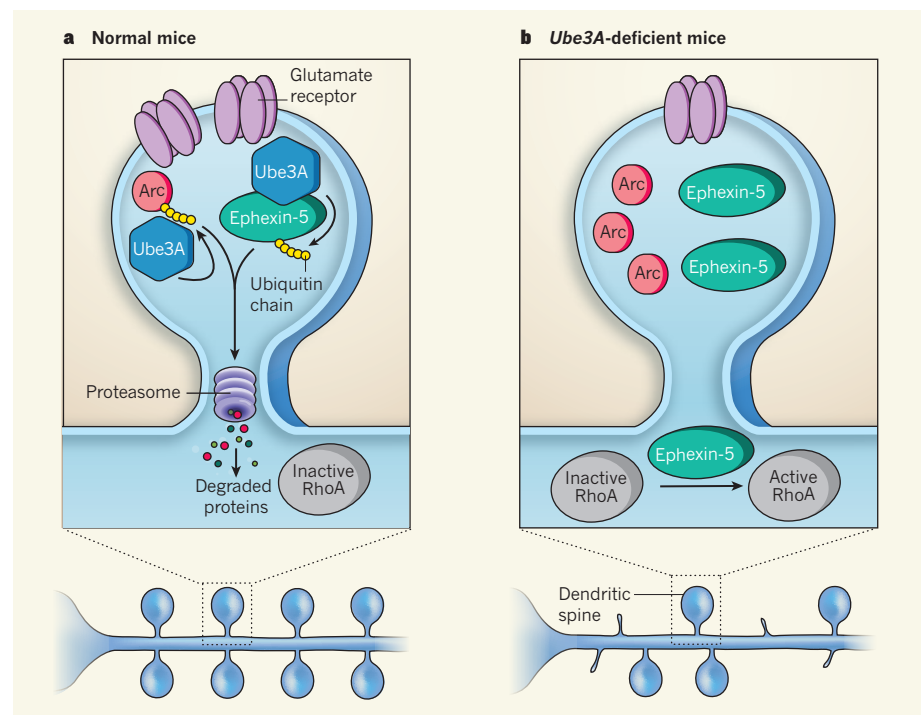
1. Bragg, L. & Nye, J. F. *Proc. R. Soc. Lond. A* **190**, 474–481 (1947).
2. Irvine, W. T. M., Vitelli, V. & Chaikin, P. M. *Nature* **468**, 947–951 (2010).
3. Van Vliet, K. J., Li, J., Zhu, T., Yip, S. & Suresh, S. *Phys. Rev. B* **67**, 104105 (2003).
4. Alsayed, A. M., Islam, M. F., Zhang, J., Collings, P. J. & Yodh, A. G. *Science* **309**, 1207–1210 (2005).
5. Yethiraj, A. & van Blaaderen, A. *Nature* **421**, 513–517 (2003).
6. Schall, P., Weitz, D. A. & Spaepen, F. *Science* **318**, 1895–1899 (2007).
7. Seung, H. S. & Nelson, D. R. *Phys. Rev. A* **38**, 1005–1018 (1988).
8. Nelson, D. R. & Peliti, L. *J. Phys.* **48**, 1085–1092 (1987).
9. Bowick, M. J. & Giomi, L. *Adv. Phys.* **58**, 449–563 (2009).

Despite remarkable progress in understanding the cellular basis of neurodevelopmental disorders, their genetic causes have proved difficult to identify. One such example is Angelman syndrome — a severe disorder that is characterized by reduced intellect, epileptic seizures, poor language skills, and unusual laughter and happiness<sup>1</sup>. Although the pathogenesis remains to be clearly defined, there is unequivocal genetic evidence that disruptions in the *UBE3A* gene can cause Angelman syndrome<sup>2</sup>. Two papers<sup>3,4</sup> in *Cell* show that, in mice, the *UBE3A* protein regulates several aspects of neuronal connectivity and function.

Genetically engineered mice lacking a functional *Ube3A* gene represent a model for

Angelman syndrome. These mice recapitulate several neurological traits observed in human patients, including learning deficits and specific alterations in neuronal circuitry<sup>5</sup>. The formation of synaptic connections between individual neurons in the brain initially appears normal. However, the refinement of synaptic connectivity is disrupted, leaving neuronal networks with reduced numbers of dendritic spines — the neuronal structures that harbour most synapses<sup>4</sup>. The neuronal networks of the mutant mice also exhibit an unusual rigidity (or lack of plasticity), which is thought to underlie the cognitive alterations in patients and their inability to learn relatively easy tasks.

These disease-related defects highlight a fundamental question of neurodevelopment: how do neurons specify and control



**Figure 1 | UBE3A and excitatory synapses.** **a**, In normal neurons, UBE3A tags substrate proteins Arc<sup>4</sup> and Ephexin-5 (ref. 3) with ubiquitin, targeting them for proteasomal degradation. Decreasing the levels of Arc and Ephexin-5 is a key determinant of dendritic-spine density, maturation and synaptic function. **b**, In *Ube3A*-deficient neurons, Arc and Ephexin-5 accumulate. High Ephexin-5 levels limit the growth and maturation of dendritic spines by promoting RhoA activity. Increased Arc levels enhance removal of the AMPA-type glutamate receptors, leading to a decrease in synaptic function.



the number of connections that they form? In other words, are there brakes that prevent premature and overabundant connectivity at early developmental stages? And how might such brakes be removed for synapse formation to proceed?

Functional analysis of UBE3A has provided some clues with which to answer these questions. This protein is an E3 ubiquitin ligase — an enzyme that transfers the small protein ubiquitin onto substrate proteins. Once tagged, these substrates are degraded by the proteasome complex and are removed from the cell. UBE3A thus controls the levels of specific substrate proteins<sup>5</sup>.

In Angelman syndrome, the neuronal level of UBE3A substrate proteins is abnormally high, causing perturbed brain development and function<sup>7</sup>. Although some UBE3A target proteins have been identified, the key substrates responsible for the development of Angelman syndrome have remained obscure. Margolis *et al.*<sup>3</sup> and Greer *et al.*<sup>4</sup> identify two UBE3A substrates that may explain several aspects of neuronal dysfunction in this disorder.

Margolis and colleagues provide compelling evidence that UBE3A ubiquitinates a neuronal signalling molecule called Ephexin-5 (Fig. 1a). Levels of Ephexin-5 are increased in mouse models of Angelman syndrome. In a series of experiments, the authors probe what happens to neuronal circuits when cells express either too much or too little Ephexin-5. They find that this protein restricts synaptic development. During early postnatal life, Ephexin-5 levels are high and keep the number of synaptic connections low, so that neurons do not indiscriminately form synapses with any cell that they contact.

Ephexin-5 blocks synapse formation by activating the signalling protein RhoA — a GTPase that limits the growth of dendritic spines. However, when two neurons interact, activation of the signalling receptor EphB promotes phosphorylation of Ephexin-5. Phosphorylated Ephexin-5 is subsequently ubiquitinated by UBE3A, tagging it for degradation. Once Ephexin-5 is removed, the establishment of new synaptic connections can proceed. However, in Angelman-syndrome neurons UBE3A is absent, allowing Ephexin-5 to accumulate; consequently, the cells exhibit increased RhoA activity and form too few synapses (Fig. 1b), resulting in dysfunction of neuronal networks.

Whereas Ephexin-5 controls synapse number, Arc — the other UBE3A target, which Greer *et al.*<sup>4</sup> identified — controls synapse function and plasticity. Specifically, Arc promotes the removal of synaptic receptors that respond to the neurotransmitter glutamate; this results in either a reduction or silencing of synaptic communication<sup>6</sup>. Greer and colleagues discovered that UBE3A binds to and ubiquitinates Arc, thereby promoting its

degradation. In UBE3A-deficient neurons, Arc accumulates, leading to a decrease in synaptic levels of glutamate receptors and a disruption of normal synaptic function (Fig. 1).

The discovery of two UBE3A substrates provides a mechanistic explanation for at least some aspects of Angelman syndrome at a cellular level. Dysregulation of Ephexin-5 and Arc represent potential causes of alterations in synapse number and plasticity, respectively, both of which are hallmarks of this disorder.

Notably, these findings may also have implications for autism. Some patients with autism carry a duplication of the genomic region 15q11–13, which contains, among other genes, *Ube3A*. This mutation results in increased expression of UBE3A and other proteins<sup>2</sup>. Extrapolating from the data of Margolis *et al.*<sup>3</sup> and Greer *et al.*<sup>4</sup>, increased UBE3A levels result in reductions in Ephexin-5 and Arc levels, thereby increasing neuronal connectivity and synaptic transmission. These hypotheses can be tested in recently generated<sup>7</sup> mouse models carrying a duplication of this genomic region.

These fresh insights<sup>4,5</sup> into neuronal dysfunction in Angelman syndrome also point to candidate drug targets. For example, some of the traits associated with the *Ube3A*

mutation in mice can be reversed by activation of calmodulin-dependent kinase 2, a signalling protein that has been implicated in promoting the delivery of glutamate receptors to synapses<sup>8</sup>. The new findings provide a first step in the attempt to alleviate the neuronal defects of Angelman syndrome, as well as related neurodevelopmental disorders such as autism. ■

**Peter Scheiffele** is in the Biozentrum of the University of Basel, 4056 Basel, Switzerland.

**Asim A. Beg** is in the Department of Pharmacology, University of Michigan, Ann Arbor, Michigan 48109, USA.  
e-mails: peter.scheiffele@unibas.ch; asimbeg@umich.edu

1. Williams, C. A. *et al.* *Am. J. Med. Genet. A* **140**, 413–418 (2006).
2. Nicholls, R. D. & Knepper, J. L. *Annu. Rev. Genomics Hum. Genet.* **2**, 153–175 (2001).
3. Margolis, S. S. *et al.* *Cell* **143**, 442–455 (2010).
4. Greer, P. L. *et al.* *Cell* **140**, 704–716 (2010).
5. Jiang, Y. *et al.* *Neuron* **21**, 799–811 (1998).
6. Chowdhury, S. *et al.* *Neuron* **52**, 445–459 (2006).
7. Nakatani, J. *et al.* *Cell* **137**, 1235–1246 (2009).
8. van Woerden, G. M. *et al.* *Nature Neurosci.* **10**, 280–282 (2007).

**P.S. declares competing financial interests.**  
See online article for details.

#### PHYSICAL CHEMISTRY

## Model's reputation restored

**The structure of a mineral has been validated, ending the controversy about its potential usefulness as a model of an unusual magnetic lattice. This model might provide insight into superconductivity.**

MARK A. DE VRIES & ANDREW HARRISON

The mineral herbertsmithite has been hailed as a rare model of an unconventional type of magnetism thought to have a key role in the mechanism of high-temperature superconductivity (a form of superconductivity that occurs above 30 kelvin). But concerns have been raised that chemical disorder in this material could produce defects in the array of magnetic atoms, disturbing or even destroying the properties that make it a useful model. In the *Journal of the American Chemical Society*, Daniel Nocera and his group<sup>1</sup> now report that they have clearly identified the type of disorder present, and have found it to have little influence on the magnetic lattice of the mineral. This represents a crucial step in establishing a simple, clean system to provide unambiguous insight into this important form of magnetism.

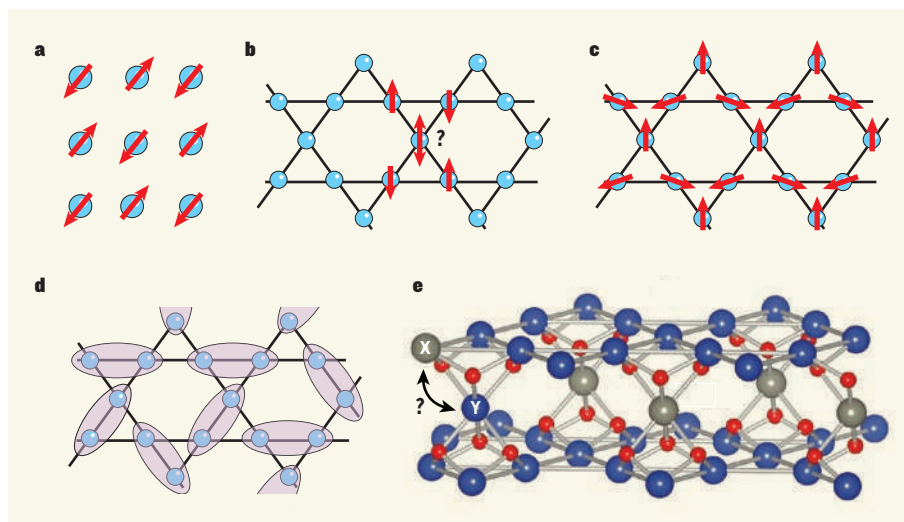
A material's magnetism is ultimately derived from unpaired electrons, each of which has a property called spin (*S*), with a value of ½, that bestows on each of them a magnetic moment. In insulators, such moments or spins are localized on atoms, and commonly interact with their closest neighbours so that specific spin orientations are preferred. In most cases, an antiparallel (antiferromagnetic) configuration of nearest-neighbour spins is favoured so that, when a lattice of such atoms is cooled, their spins usually freeze to form an ordered array (Fig. 1a). However, under certain circumstances — for some magnetic lattices, for instance — quite different behaviour can ensue. One such case is the kagome antiferromagnet, which is formed from corner-sharing triangles (Fig. 1b). In these systems, it is impossible to arrange each near-neighbour pair of spins so that they are all antiparallel, and the system is said to be geometrically frustrated.

The lowest-energy arrangement of a kagome antiferromagnet might be expected to be one in which neighbouring spins are oriented  $120^\circ$  to each other (Fig. 1c). However, the physicist Philip Anderson pointed out that, for several lattices composed of triangles<sup>2</sup> in which there is just one electron per atom (that is, when  $S = \frac{1}{2}$ ), it is more favourable for neighbouring spins to pair up in a manner analogous to that of the electron spins in chemical bonds. Quantum mechanics allows the spins in 'valence' bonds to be simultaneously up–down and down–up (where 'up' and 'down' are the two possible alignments of the spins), thus relieving geometric frustration by effectively allowing pairs of local moments to cancel out.

Because such coupling may occur between all pairs of spins, the overall picture is that of a 'liquid' of valence bonds, a state that resonates between all possible ways of making such bonds (Fig. 1d). This state is called a quantum spin liquid, or a resonating valence bond (RVB) liquid (by analogy with Linus Pauling's model<sup>3</sup> of chemical bonding in organic molecules such as benzene, in which the chemical structure can also be described as a hybrid of different valence-bond arrangements). It was proposed<sup>4</sup> — again by Anderson — that in the cuprate class of high-temperature superconductors, the RVB state enables the formation of superconducting charge carriers.

Among lattices of triangles, kagome lattices that have antiferromagnetically coupled spins of  $S = \frac{1}{2}$  have been regarded for some time as prime candidates for a RVB state. But an undistorted realization of this highly prized system — herbertsmithite, which has the nominal formula  $\text{ZnCu}_3(\text{OH})_6\text{Cl}_2$  — was only recently reported<sup>5</sup> by Nocera's group. To be precise, it is the array of spins on the copper ions ( $\text{Cu}^{2+}$ ) in herbertsmithite that could form an RVB magnetic state (Fig. 1e).

It was soon demonstrated<sup>6</sup> that, although there is a very strong antiferromagnetic coupling between spins in the mineral, no spin freezing could be observed, even at temperatures as low as 50 millikelvin. This is consistent with the existence of some form of RVB state in herbertsmithite. But evidence was also found for disorder in the compound's structure, raising concerns that it was not as clean a model system as had been hoped. Neutron-diffraction data, combined with elemental analysis using inductively coupled plasma Auger electron spectroscopy (ICP-AES) on herbertsmithite indicated that about one-quarter of the sites expected to be occupied by zinc ions ( $\text{Zn}^{2+}$ ) are actually occupied by  $\text{Cu}^{2+}$  (Fig. 1e), and that one-twelfth of the  $\text{Cu}^{2+}$  sites on the kagome lattice are occupied by  $\text{Zn}^{2+}$ . The actual level of disorder was estimated to be slightly lower, however, on the basis of analyses of the compound's magnetic susceptibility<sup>7</sup> and heat capacity<sup>8</sup>. Both of these properties contain a contribution from nearly free spins (which interact only very weakly with other



**Figure 1 | Spin arrangements in crystal lattices.** Unpaired electrons on atoms have a magnetic moment, or spin. These spins adopt preferred alignments (indicated by red arrows) in crystal lattices. **a**, Antiferromagnetic alignments, in which all neighbouring spins are antiparallel (as in this square lattice) are often favoured. **b**, In kagome lattices, triangles of atoms are joined at their corners, making a completely antiparallel spin arrangement impossible. **c**, This can lead to a compromise arrangement in which the spins are oriented at  $120^\circ$  to each other. **d**, Alternatively, a quantum spin liquid might form as a combination of many states (one of which is depicted here) in which spins pair up, analogous to the pairing of electrons in chemical bonds. **e**, The mineral herbertsmithite ( $\text{ZnCu}_3(\text{OH})_6\text{Cl}_2$ ) contains kagome layers of  $\text{Cu}^{2+}$  ions (blue) linked by  $\text{O}^{2-}$  ions (red), separated by layers of  $\text{Zn}^{2+}$  (grey) and  $\text{Cl}^-$  ions (not shown). Defects occur in the lattice when  $\text{Cu}^{2+}$  ions occupy  $\text{Zn}^{2+}$  sites and vice versa (indicated by the double-headed arrow between atoms marked 'X' and 'Y'). Nocera and colleagues<sup>5</sup> find that the number of defects in herbertsmithite is smaller than had been thought. This makes the mineral an ideal model for studying spin-liquid states, which have been implicated in high-temperature superconductivity.

spins in the material). This contribution can be attributed to the  $S = \frac{1}{2}$  spins of  $\text{Cu}^{2+}$  ions occupying one-fifth of the  $\text{Zn}^{2+}$  sites.

Nocera and colleagues appreciated that the uncertainty about the disorder was unsatisfactory — they felt that rigorous determination of the exact level of disorder should depend on structural and chemical methods alone. They therefore performed<sup>1</sup> new measurements on herbertsmithite, using a suite of techniques: X-ray absorption spectroscopy and advanced X-ray- and neutron-diffraction methods. Taken together, their results show that about 15% of the  $\text{Zn}^{2+}$  sites are occupied by a  $\text{Cu}^{2+}$  ion, thus confirming the earlier interpretations of the magnetic susceptibility<sup>7</sup> and heat capacity<sup>8</sup> of herbertsmithite, and convincingly dispelling the last doubts about the concentration of  $\text{Cu}^{2+}$  ions (and so of  $S = \frac{1}{2}$  spins) on the  $\text{Zn}^{2+}$  sites. These defect spins will interact only weakly with other spins, and thus will have very little influence on the behaviour of the spins on the kagome lattice.

What comes as a surprise in Nocera and colleagues' results<sup>1</sup> is that the concentration of  $\text{Zn}^{2+}$  ions in the kagome  $\text{Cu}^{2+}$  sites was found to be only about  $1 \pm 3\%$ . This is good news for those investigating quantum magnetism — it means that there are very few vacancies in the periodic array of spins forming the kagome lattice, allowing for a reliable comparison between theory and experiment in herbertsmithite. The lower than expected chemical disorder implies that the chemical formula of herbertsmithite

is probably closer to  $\text{Zn}_{0.85}\text{Cu}_{3.15}(\text{OH})_6\text{Cl}_2$ , rather than  $\text{ZnCu}_3(\text{OH})_6\text{Cl}_2$ , as was thought previously. This revised formula will doubtless be checked in other laboratories using the ICP-AES method of chemical analysis.

As the first RVB system in which spin correlations and dynamics can readily be measured, herbertsmithite has great potential to reveal the character of spin liquids. More broadly, it should also allow an exploration of the relationship between antiferromagnetism and superconductivity in layered transition-metal compounds. ■

**Mark A. de Vries** is at the School of Physics and Astronomy, University of Leeds, Leeds LS2 9JT, UK. **Andrew Harrison** is at the Institut Laue-Langevin, 6 rue Jules Horowitz, F-38042 Grenoble, France, and at the School of Chemistry, University of Edinburgh, Edinburgh EH9 3JJ, UK. e-mails: m.a.devries@leeds.ac.uk; harrison@ill.eu

1. Freedman, D. E. et al. *J. Am. Chem. Soc.* **132**, 16185–16190 (2010).
2. Anderson, P. W. *Mater. Res. Bull.* **8**, 153–160 (1973).
3. Pauling, L. *The Nature of the Chemical Bond* 3rd edn (Cornell Univ. Press, 1960).
4. Anderson, P. W. *Science* **235**, 1196–1198 (1987).
5. Shores, M. P., Nytko, E. A., Bartlett, B. M. & Nocera, D. G. *J. Am. Chem. Soc.* **127**, 13462–13463 (2005).
6. Mendels, P. et al. *Phys. Rev. Lett.* **98**, 077204 (2007).
7. Bert, F. et al. *Phys. Rev. B* **76**, 132411 (2007).
8. de Vries, M. A., Kamenev, K. V., Kockelmann, W. A., Sanchez-Benitez, J. & Harrison, A. *Phys. Rev. Lett.* **100**, 157205 (2008).



# Emerging properties of animal gene regulatory networks

Eric H. Davidson<sup>1</sup>

Gene regulatory networks (GRNs) provide system level explanations of developmental and physiological functions in the terms of the genomic regulatory code. Depending on their developmental functions, GRNs differ in their degree of hierarchy, and also in the types of modular sub-circuit of which they are composed, although there is a commonly employed sub-circuit repertoire. Mathematical modelling of some types of GRN sub-circuit has deepened biological understanding of the functions they mediate. The structural organization of various kinds of GRN reflects their roles in the life process, and causally illuminates both developmental and evolutionary process.

The body plan of an animal, and hence its exact mode of development, is a property of its species and is thus encoded in the genome. Embryonic development is an enormous informational transaction, in which DNA sequence data generate and guide the system-wide spatial deployment of specific cellular functions. GRNs also determine the main events of postembryonic development, including organogenesis and formation of adult parts and cell types. Beyond that, GRNs control a vast array of physiological capabilities and modes of response to environmental fluctuations and challenges. GRNs are composed of multiple sub-circuits, that is, the individual regulatory tasks into which a process can be parsed are each accomplished by a given GRN sub-circuit<sup>1–4</sup>. Thus the operational significance of a GRN structure will be indicated by the types of sub-circuit it contains. However, GRNs have more global organizational properties as well. The comparative review below shows that GRNs may be deeply layered, generating successive regulatory transactions, or they may be shallow, in the sense that they mandate few transactions between the initial inputs and the terminal activation of effector genes.

## The developmental GRN sub-circuit repertoire

Modular GRN sub-circuits are defined by their topologies, and the topology of a sub-circuit directly indicates its function in life. In this article I am concerned only with sub-circuits which perform developmental biology jobs that can be defined uniquely, and not with very common 'motifs' such as the coherent feed forward loop, which although it has specific dynamic properties<sup>5</sup>, appears in so many different contexts that no unique developmental biology function can be associated with it. Table 1 contains a compilation of sub-circuits drawn from all the various GRNs considered in this review, together with an abbreviated description of their regulatory functions, and abbreviated diagrams illustrating the canonical sub-circuit structures. Additional sub-circuits will be found as more developmental GRNs are explored, but the basic import of Table 1 is that there probably exists a small, finite number of sub-circuit topologies out of which developmental programs of all kinds are constructed. The first entry, for example, is a spatial information processing sub-circuit called the double-negative gate, found in the sea urchin embryo GRNs<sup>2,6</sup>. This sub-circuit consists of two genes encoding repressors wired in tandem, so that the target of the first repressor is the gene encoding the second, plus downstream regulatory genes which are targets of the second repressor. Its function is to ensure that the target genes are expressed only where the first repressor

is (transiently) active (domain X), while these genes are shut down everywhere else ( $1 - X$ ); what we term an  $X, 1 - X$  processor.

References in Table 1 generalize the point that structurally similar sub-circuits, but composed of different regulatory genes, are repeatedly encountered doing similar developmental jobs in different GRNs. At root this is because what the circuit can do depends directly on its structure; for example, in a recent study, a search of all possible small sub-circuits based on 3-node topologies showed that only two are capable of response to a signal followed by return to the original state<sup>7</sup>.

A given sub-circuit structure implies a given function, and in development there is a finite set of regulatory functions required. The biochemical complexity of the diverse *cis*-regulatory systems composing developmental GRN sub-circuits, and the diversity of the sets of transcription factors which animate them, thus may give way to a pleasingly simple set of logic-processing sub-circuit topologies. This would be a very important outcome, for it would make it possible to parse the apparently enormous mazes of interconnections in system level GRNs into modules of developmental logic function, and thus to understand how GRNs control the biology we see.

## Deep structure of embryonic GRNs

The GRNs which control the *de novo* formation of embryonic territories typically include many different functional sub-circuits, which govern successive 'layers' of process. They are hierarchical in their overall structure. Their depth simply reflects the long sequence of regulatory steps required to complete any component of embryonic development. The concept of deep as opposed to shallow GRN structure can be simply considered as the number of successive changes in regulatory state required to generate an episode of embryological or other development, between the initial state and the terminal process which the GRN causes to happen. The terminal outcome is, by definition, the activation of cohorts of effector genes (that is, differentiation and cell biology genes, as opposed to only regulatory genes). In relatively shallow GRNs, some of which are considered below, the initial state may be a paused regulatory condition just upstream of expression of a differentiation gene battery.

The sea urchin embryo endomesoderm GRN serves as a reference point, as at present it is the most nearly complete, predictively useful, and validated large scale developmental GRN available. Structure/function aspects of this GRN have been reviewed recently<sup>2</sup>, and an always current version, together with underlying data and dynamic presentations by domain, is available at <http://sugp.caltech.edu/endomes/>. The

<sup>1</sup>Division of Biology 156-29, California Institute of Technology, Pasadena, California 91125, USA.

**Table 1 | Sub-circuit repertoire for developmental GRNs**

Regulatory state specification function	Sub-circuits	What they do	Topologies
X, 1 – X processors	Double negative gate <sup>1,2,6</sup> 1.1	Install regulatory state in X domain, prohibit same state everywhere else*	
	Signal-mediated switch <sup>2</sup> 1.2	Activate regulatory gene(s) in cells receiving signal, repress same genes everywhere else†	
Spatial subdivision	Inductive signaling <sup>2</sup> 2.1	Activation of new regulatory genes in a cellular domain by transcriptional response to signal ligands produced by other cells	
	AND logic circuitry <sup>2</sup> 2.2	Overlapping but spatially non-coincident inputs are generated and both are required for regulatory gene activation, which occurs only in overlap subdomain	
	Spatial repression <sup>2</sup> 2.3	Boundaries of spatial regulatory state domains controlled by transcriptional repression.	
Dynamic lockdown of regulatory state‡	Reciprocal repression of state <sup>1,2,9,39,51</sup> 3.1	In each spatial regulatory state domain key activators of alternative states are transcriptionally repressed by 'exclusion' circuitry§.	
	Feedback circuitry <sup>1</sup> 3.2	Two or three regulatory genes engage in positive intergenic feedback, stabilizing regulatory state irrespective of transient inputs	
	Community effect circuitry <sup>2,23,92</sup> 3.3	Cells within a territory all signal to one another, driving continued uniform expression both of ligand gene and signal-dependent regulatory genes	

concept of GRN depth is illustrated in Fig. 1a by abstracting from the sea urchin GRN the sequence of sub-circuits deployed in order to specify its skeletogenic cell lineage<sup>1</sup>, which produces only the one cell type, and is developmentally the simplest process modelled in the whole endomesoderm GRN. This portion of the network contains 24 regulatory genes and several signalling genes, as well as a sampling of downstream differentiation genes. Without detailing the individual genes and linkages in the skeletogenic GRN, its internal structure is abstractly represented in Fig. 1a as a series of interconnected boxes, each of which represents a GRN sub-circuit that executes the indicated regulatory task. Many of these sub-circuits are among the types listed in Table 1, as indicated by the colour coding, and the arrows show the linkages from one sub-circuit to another, that is, they represent transcription factors generated in one box and used for control of gene(s) in the next box, the inputs or 'feeds' into each sub-circuit. The boxes are layered hierarchically, with those that initiate the process at the top. Figure 1a includes various control processes that are common throughout embryonic development, because the problems that have to be solved are general: the initial spatial inputs have to be interpreted, the regulatory state then has to be

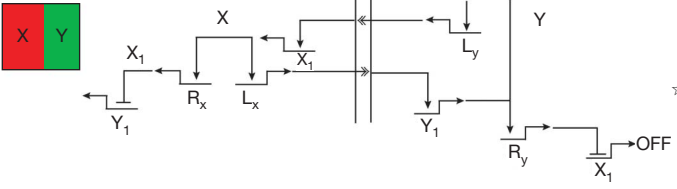
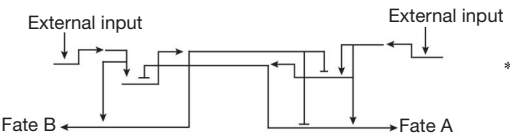
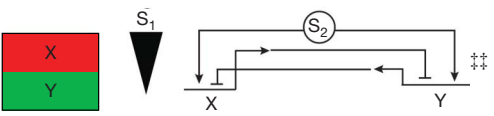
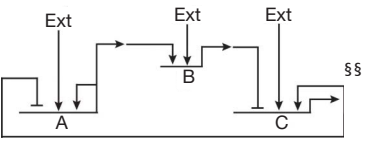
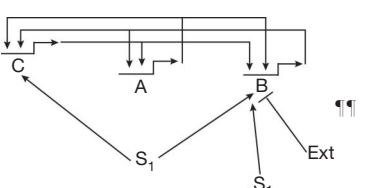
locked down (the initial inputs are always transient), signals have then to be generated, other states have to be excluded, and differentiation drivers have to be activated. It is not surprising that all this requires a lot of sequential circuitry, even given the relative simplicity of skeletogenic lineage development. The GRNs underlying specification of the mesoderm, endoderm<sup>1,2,8</sup> and of the oral and aboral ectoderm<sup>9</sup> of the sea urchin embryo are similarly deep, layered and hierarchical.

GRNs for specification of mesoderm in *Xenopus* embryos<sup>10</sup>, for specification of the gut and mesoderm cell lineages in *Caenorhabditis elegans* embryos<sup>11</sup>, and for specification of endoderm and dorsal, anterior and ventral gene expression domains<sup>12</sup>, and of mesoderm<sup>13</sup> in zebrafish, also display deep, hierarchical organizations. Some of these GRNs are compilations from the literature or from chromatin immunoprecipitation (ChIP)-chip observations, and have not been validated by direct perturbation analysis, let alone at the *cis*-regulatory level, but it is unlikely that their overall structure is illusory.

The GRN for dorsal/ventral patterning in *Drosophila*<sup>14</sup>, which does have extensive *cis*-regulatory support, is also hierarchical, but its unusual structure reflects the unusual developmental process it controls. In this



Table 1 | Continued

Regulatory state specification function	Sub-circuits	What they do	Topologies
Boundary maintenance	Reciprocal repressive signalling across boundary <sup>32</sup>	Different signals are produced by apposing cells and their reception triggers repressive circuitry excluding the cross-boundary regulatory state	
Terminal binary cell fate choice	Alternate sub-circuits driven by reciprocal repressors <sup>8,36-38,81</sup>	External inputs tip the balance of repressor expression, resulting in activation of one differentiation program and exclusion of the other	
Discontinuous transcriptional response to signal intensity and/or duration	Reciprocal repressor genes responding cooperatively to inducer <sup>21,38,81</sup>	Circuitry generates differential stimulation of expression of reciprocal repressors in low versus high signal intensity††	
	Reciprocal repressor genes, one activating an additional repressor gene, each with variable external positive inputs <sup>82</sup>	Circuitry generates irreversible transitions, in stem cell regulatory state, off versus on in response to signals of different strength and duration	
	Triple feedback linkage with asymmetric signal inputs <sup>47</sup>	Produces alternative regulatory states, or low level indeterminate state, depending of different positive inputs	

The role of the sub-circuit is given in column 1; its name in column 2; a description of its function in column 3; and the sub-circuit structure in column 4. Numbers in column 2 are keyed to Fig. 1. See references indicated for actual occurrences, exact circuit topologies, and discussion of information processing specifics. In each case the functions of the circuit are hardwired in its *cis*-regulatory target sites. In Topologies column, all genes encode transcription factors unless otherwise noted.

\* Regulatory genes that create initial regulatory state are controlled by widely expressed repressor, which is dominant over their positive inputs, and gene encoding this repressor is itself specifically repressed in a local region (X) by another gene encoding a different repressor: hence target genes are ON in X, specifically repressed elsewhere.

† Many developmental signalling systems (for example, Notch, Wnt) activate immediate early response factors in cells receiving ligand, but in absence of ligand, these factors act as dominant repressors of the same target genes.

‡ Dynamic in that continuing transcription is required.

§ Exclusion sub-circuits are activated as downstream outputs of specification GRNs.

|| A unique circuit design here is that the ligand gene is activated by the same signal transduction mechanism reception of the ligand activates in recipient cells; a positive intercellular feedback.

¶ From ref 2.

# L, gene encoding signalling ligand.

☆ R encodes repressor; L encodes signalling ligand.

\*\* This example was adapted from Ref 36.

†† Conceived as a means of obtaining different discrete transcriptional responses from a graded signal; see discussion of this type of circuitry in section on mathematical models below.

‡‡ S, signal; triangle represents graded signal strength

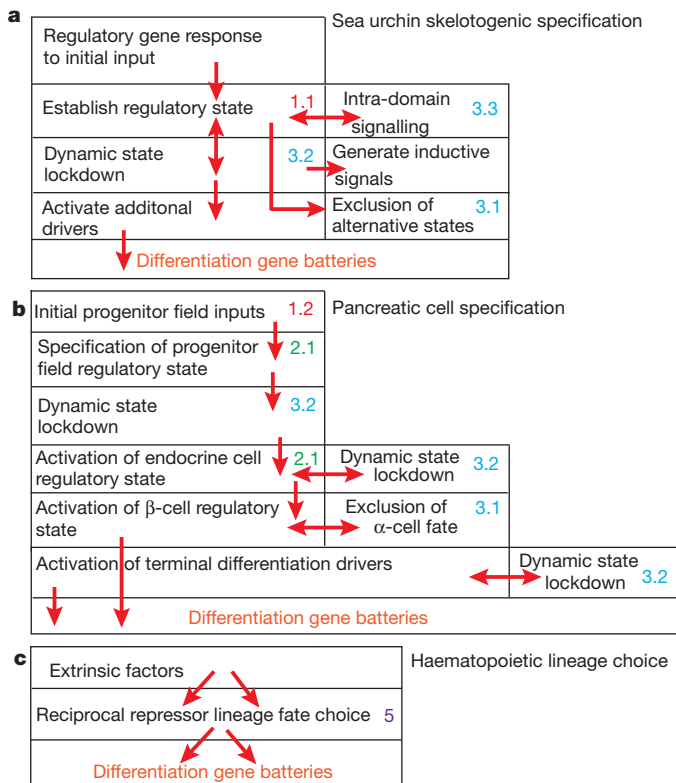
§§ S<sub>1</sub>, S<sub>2</sub>, different signal inputs gene B is subject to additional transcriptional repression in certain regulatory states.

|||| This design precludes necessity for ad hoc Hill coefficients as in 5, 6.1; see section on mathematical models below.

¶¶ Autoregulatory loops lock on whichever state the system goes to.

embryo regulatory state domains are initially set up very quickly in a syncytium, without intercellular signalling, although following cellularization, signalling dominates further transcriptional functions and the GRN henceforth has a typical structure. In the syncytial embryo spatial stripes of both dorsal/ventral and anterior/posterior regulatory gene expression, which specify the future multicellular embryonic territories, are generated by parallel *cis*-regulatory responses to maternally localized and zygotically expressed combinations of diffusing transcription factors<sup>15-17</sup>. Many of these factors act as repressors in setting spatial boundaries. Important initial inputs in this system are the transcription factors Dorsal and Bicoid, encoded by maternal messenger RNAs which become distributed in graded fashion in the syncytial embryo nuclei, from ventral to dorsal and anterior to posterior, respectively. *Cis*-regulatory modules have been isolated that control target genes expressed in stripes at given ranges of values of these 'morphogens'. When associated with reporters, and introduced into the egg, these *cis*-regulatory

modules produce stripes of expression at the appropriate positions along the respective axes. It has been assumed for a long time that the positions where they operate are determined by the quantitative values of the Dorsal or Bicoid concentrations at those locations, which in some way these *cis*-regulatory systems read. Much recent evidence, however, shows that the positions where these *cis*-regulatory modules act depend on combinatorial activator and repressor inputs<sup>15-19</sup>, and the quantitative values of these 'morphogens' alone do not by themselves predict the spatial expression domains of their target genes (except sometimes at the extreme positions where their concentrations are highest). The combinatorial inputs are the products of regulatory genes linked into the GRN<sup>20</sup>. Thus the concept that the position of target gene expression is determined solely by the quantitative value of the 'morphogen' is overly simplistic: the overall pattern, and the overall signal strength response mechanism, are actually network properties rather than a property of individual *cis*-regulatory modules that independently and quantitatively read single gradient values.



**Figure 1** | ‘Birdseye’ views of structural properties of representative developmental GRNs. **a–c**, Diagrammatic view of sub-circuits and sub-circuit functions in three different GRNs. Each box represents a GRN sub-circuit consisting of a small number of regulatory genes and their functional linkages. Coloured dots and numbers refer to the similarly coded sub-circuit types in Table 1. Red arrows indicate linkages between sub-circuits, that is, regulatory feeds from one sub-circuit to another. **a**, GRN for skeletogenic mesoderm lineage specification in sea urchin embryos<sup>1</sup>. **b**, GRN for pancreatic developmental process<sup>3</sup>, leading to  $\beta$  cell specification and insulin gene transcription. **c**, GRNs typical of terminal binary fate choices in haematopoietic stem cells and other similar situations, as discussed in text.

Vertebrate embryos also use graded inputs, but in this case the developmental systems are cellular, and the ‘morphogens’ are diffusible extracellular signal ligands. Distinct regulatory responses occur, dependent on the intensity of signalling, resulting in activation of different genes in different locations in the embryo, for example in response to an activin gradient in the pre-gastrular *Xenopus* embryo<sup>21</sup>. Modelling shows that this particular level-sensitive response could be mediated by a specific type of GRN sub-circuit, in which regulatory genes encoding repressors reciprocally damp each other’s expression, while responding differentially in a cooperative, thus nonlinear and discontinuous way, to the concentration of the signal (Table 1). However, as we see in the following there is more than one type of sub-circuit capable of discontinuous response to a graded signal.

## Structure of GRNs encoding body parts

We now have bits and pieces of the GRNs controlling development of body parts and organs in later embryonic development, usually their initial stages. Like embryonic GRNs, they are deep and hierarchical, and indeed are in most ways structured similarly to the early embryo GRNs. The box diagram cartoon in Fig. 1b provides an example. This diagram is abstracted from a GRN for specification of pancreas and then of pancreatic  $\beta$ -cells<sup>3</sup>. In adult body part formation, including organogenesis, the first step is always establishment of a given regulatory state in the field of cells from which the body part will form, the progenitor field<sup>22</sup>, for example, the cardiac crescent, or the limb bud, or the imaginal disc. The progenitor field is positioned with respect to the coordinates of the developing organism, which always involves signal-mediated

installation of a new regulatory state. But then the field is subdivided into the regulatory state domains of its subparts, and at each step the state is locked down. This used to be called ‘pattern formation’, when people were looking at only one or a few genes at a time. As in early embryo GRNs the main job in setting up the parts and future form of the organ is the progressive deployment of regulatory states in space. It is essential to realize that this process is not to be equated with terminal cell fate specification; the cells expressing patterned regulatory states are yet far upstream in the developmental process from their ultimate descendants, which will eventually differentiate in various directions, according to what part of the organ they arise in.

The similarity between GRNs encoding adult body part formation and those controlling earlier embryogenesis is also sustained at the sub-circuit level in that few additional types of sub-circuit are used. For instance, in pre-gastrular embryonic specification it can be confidently predicted that a feedback circuit locking two or three regulatory genes in a mutual positive embrace will be encountered just downstream of the initial inputs used to set up a given regulatory state (Table 1). This is seen at multiple locations in the sea urchin embryonic GRNs<sup>1–3,23</sup>, and the same feature routinely appears in adult body part GRNs: for example, in those underlying development of neural crest in lamprey<sup>24</sup>, gut specification in vertebrates<sup>25</sup>, eye lens field specification in both vertebrates<sup>26</sup> and *Drosophila*<sup>27</sup>, haemangioblast specification in mouse<sup>28</sup>, pharynx specification in *C. elegans*<sup>29</sup>, heart specification in mammals and *Drosophila*<sup>3,30,31</sup>, and pancreatic  $\beta$ -cell specification in mouse<sup>3</sup>. Each of these GRNs include two- or three-gene positive feedback sub-circuits functioning to lock down newly installed spatial regulatory states. The similar feedback sub-circuits are constructed with different genes; again it is the sub-circuit topology that determines function, and many different regulatory genes can have the same roles. An additional type of sub-circuit that is commonly seen in later embryonic processes is a signal-mediated, mutual repression device that operates across a cellular boundary, such that, on either side, reception of a signal from the other side specifically causes repression of key genes of the alternate regulatory state (Table 1). Four examples from later development where this type of circuitry obtains are in the GRN that maintains the distinct regulatory states of anterior and posterior parasegment compartments in *Drosophila*<sup>32</sup>, the GRN controlling establishment of dorsal/ventral neural tube domains in vertebrates<sup>33</sup>, the GRN controlling anterior versus posterior specification in the vertebrate limb bud<sup>34</sup>, and the GRN controlling cell type specification under signal control in the *C. elegans* vulva<sup>35</sup>.

## Postembryonic developmental GRNs: differentiation from pluripotent stem cells

A remarkably recurrent similarity in GRN circuit design has recently emerged in studies of the transcriptional pathways that control binary fate choices executed in the diversification of haematopoietic cell types from multipotent precursors (for reviews, see refs 36–38). At the cores of these circuits, which use some overlapping and some lineage-specific regulatory genes, are pairs of genes encoding transcription factors that mutually antagonize each other’s expression within the same nucleus. Often initially co-expressed at relatively low levels, the lineage fate choice depends on stepped up asymmetric expression of one or the other of the core repressor gene pair. Each of these genes also directly or indirectly promotes expression of positive regulators necessary for execution of one of the lineage fate choices. As the activity of one of the core repressors increases, it causes transcriptional extinction of expression of the alternative choice, and the irreversible installation of its own positive regulatory state (see discussion of the mathematical features of such circuits below). An important point is that the genes of the antagonistic repressor pairs, and/or the regulatory genes that are their immediate targets, also provide direct positive or negative inputs into terminal differentiation genes of the alternate lineages<sup>37,39</sup>. In other words, this apparatus is deployed immediately upstream of the drivers of the effector genes that generate the features of given cell types (Fig. 1c). In comparison to the embryonic GRNs just considered, these are relatively



shallow networks. Ultimately the decisive inputs into one or the other of the core repressors are provided by extrinsic signalling ligands, for example cytokines and growth factors, including Notch and Tgf $\beta$ , or endogenous immune receptor signals. The binary choice transcriptional apparatus responds to signal intensity, so that a low input gives one result and a high input another. Different pairs of repressor genes perform similar roles in different lineage fate choices, but what is remarkable is the similar circuitry adduced throughout haematopoietic diversification. Transcriptional balance between pairs of cross-antagonistic repressors decides the outcome, for instance, in myeloid progenitors giving rise to macrophages or neutrophils<sup>37</sup>; in precursors that may give rise to either B cells or macrophages<sup>38</sup>, where there is *cis*-regulatory evidence of the transcriptional cross-repression; in the upper level decision point where erythroid versus myeloid fates bifurcate<sup>40–42</sup>; in the erythroid versus platelet fate decision<sup>43</sup>. Similarly, in T-cell diversification between helper vs killer fate<sup>44,45</sup>, T-cell receptor signal strength indirectly controls repressor function, a compelling case because there is direct *cis*-regulatory evidence of the reciprocal transcriptional silencing interactions<sup>46</sup>.

Although to some it is tempting to view all development through the same lens, there are fundamental differences between the terminal fate choice circuitry discussed here and the GRNs that execute early and mid-stage embryonic development of animal body parts. Differentiation gene batteries can be activated only at the end of the series of GRN transactions that decide exactly where they are to be deployed. Haematopoietic cell fate decisions occur at the end of a complex prior developmental process, and in fact as discussed below, the circuitry controlling very early haematopoietic stem cell pluripotentiality operates in an entirely different manner from the binary choice circuitry just considered<sup>28,47</sup>. In their function, haematopoietic binary choice sub-circuits are similar to the terminal sub-circuits that elsewhere in development immediately determine deployment of differentiation gene batteries. This perhaps explains why a characteristic of the stem cell differentiation choice systems, in other words the simultaneous low level expression in the multipotent precursors of differentiation genes indicative of multiple possible fates<sup>48,49</sup> ('lineage priming'), is not seen in embryonic fate choices. That is, in embryonic body part development the spatial fate decision is made far up in the GRN hierarchy, and locked down, long before the differentiation gene battery is deployed. In contrast, in the production of functional immune cell types the last steps in the decision have to be deferred until the multipotential cells can be told which of its potentialities is more needed. Similar binary choice circuitry is also used in non-haematopoietic developmental contexts, but again at late stages in a given process where a terminal fate choice is to be made. For example, after mammalian somites have formed, they generate spatially confined subdomains, one of which is the dermomyotome. This consists of multipotent stem-like cells, where the choice to generate vascular muscle versus smooth muscle cell types is controlled by a modulated signal, mutual repression between the *Pax3* and *Foxc2* genes<sup>50</sup>. Another non-haematopoietic circuit that in essence is remarkably similar to the antagonistic haematopoietic repressor pair sub-circuits was discovered in *C. elegans*, also operating at the terminus of much prior development<sup>51</sup>. This circuit maintains the expression of distinct sets of differentiation genes expressed in left versus right taste neurons, but the duelling repressors expressed alternately in these two neurons are in this case microRNAs that directly target the mRNAs encoding the alternate differentiation drivers. All of these kinds of sub-circuits, operate to choose, and/or to maintain the choice, of one of an alternative pair of differentiation gene driver sets.

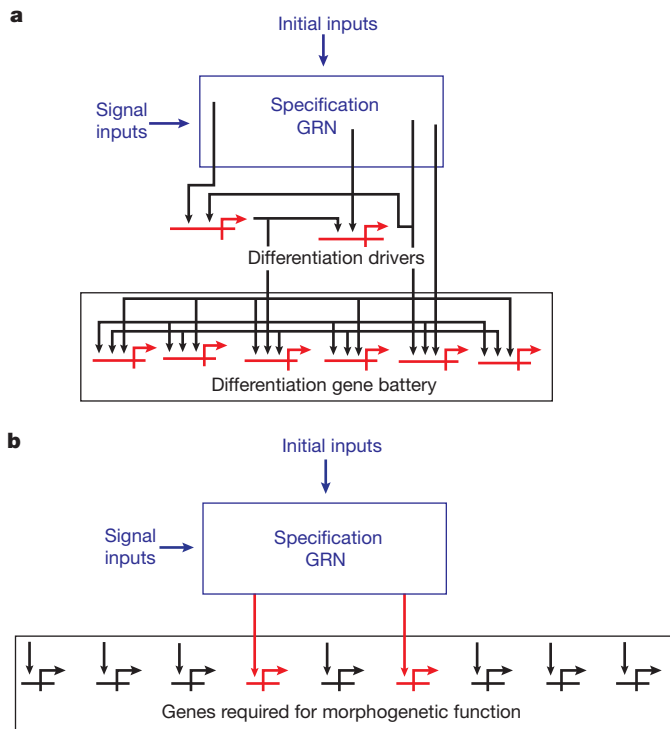
A priori, development of the body plan cannot be reduced to differentiated cell type specification, the last step in the process, nor to binary decisions between alternative fates. This is at root because development of the body plan requires a long sequence of multidimensional spatial decisions: during pattern formation spatial regulatory states must be installed progressively within multiple (>2) diverse boundaries, and also in certain anterior-posterior and dorsal-ventral positions with

respect to the body plan. In each structure of the body regulatory states that include differentiation gene battery drivers are finally installed. Thus it is not in principle surprising that if the set of differentiation gene battery regulators is changed by experimental intervention, a different cell type can be made to appear. Many recent studies show that insertion of vectors expressing sets of transcription factors or even single transcription factors can result in the change of differentiated state from one haematopoietic cell type to another<sup>36</sup>; from fibroblast to neuron<sup>52</sup>, from exocrine to pancreatic  $\beta$  cell<sup>53</sup>, etc. These cell fate changes all occur near the far downstream periphery of GRN hierarchy, as symbolized in Fig. 1. Growing a new cell type requires activation of a new differentiation battery, whereas growing a new body part requires a prior process of spatial pattern formation driven by a deep GRN. More generally, although there are embryonic processes that look superficially like the binary choices just discussed, they are effected very differently. As an example, in the sea urchin embryo, endomesodermal precursor cells give rise both to mesoderm and to endoderm, fates driven by entirely distinct regulatory states. But a careful experimental analysis<sup>8</sup> shows that there is no pluripotent 'endomesodermal' GRN, and instead a Delta/Notch signal activates a set of regulatory genes which constitute a mesoderm GRN, while in the same cells a Wnt/Tcf signal activates a different set of regulatory genes which constitute the endoderm GRN. The genes of the mesoderm GRN and of the endoderm GRN are expressed independently of one another, without any interactions. The cells of each regulatory state are then separated physically by a cell division, so that the Notch signal is received exclusively by one ring of cells, which becomes mesoderm, while the other cells express the endoderm GRN exclusively<sup>8</sup>. Nor are the exclusion functions (Table 1) that in given regulatory states act to repress genes key to alternative regulatory states 'bipotential switches'. These sub-circuits are used to lock down regulatory choices already installed rather than to make choices. They may look superficially like the mutual repression sub-circuits that switch lineages bipotentially, but they are not.

## Differentiation gene battery structure

Differentiation gene batteries account for functional cell type specificity, and a canonical network structure can be associated with them. This structure describes the topology of the regulatory relationships causing the protein coding differentiation genes of the battery to be expressed more or less coordinately. Differentiation gene batteries are per se shallow, relatively simply constructed types of sub-circuit, often wired in coherent feed forward format, as for example in sea urchin embryos<sup>1</sup>, pancreatic  $\beta$ -cells<sup>3</sup>, and macrophages<sup>54</sup>. As the immediately upstream GRNs are being uncovered, an additional characteristic of differentiation gene battery regulatory circuitry is often encountered: this is the occurrence of feedback between the drivers of the differentiation genes just upstream of the linkages to the effector genes, either auto- or cross-regulatory<sup>55,56</sup>, though this is not always seen<sup>26</sup>. The canonical form is that of Fig. 2a. Differentiation gene batteries consist of a sometimes very large number of effector genes, the relevant *cis*-regulatory modules of which (per battery) respond to members of a small set of transcription factors present as part of the terminal regulatory state. However, each such *cis*-regulatory module may in addition be serviced by some additional factors, which accounts for the fact that all the genes of the battery are not exactly expressed in lockstep<sup>3</sup>. For example, muscle protein genes are activated by two or three of the transcription factors orthologous to Srf, Mef2, and a myogenic bHLH factor in vertebrates<sup>57</sup>, plus, individually, other factors; whereas in *C. elegans* the differentiation genes of each class of neuron are identified by their response to a single key transcription factor, sometimes together with other factors<sup>55</sup>.

It is logically consistent that where there is direct repression of differentiation gene batteries by a proximal control circuit ('anti-differentiation') much the same architecture would be employed. In embryonic stem cells a hierarchical GRN that maintains the pluripotent state is headed by a recursive triple feedback system that links *Nanog*, *Oct4* (also known as *Pou5f1*) and *Sox2* genes<sup>58,59</sup>. Apparently directly downstream of this are



**Figure 2 | Structural characteristics of downstream effector gene cassettes and their control functions.** **a**, Typical differentiation gene battery, as discussed elsewhere<sup>3</sup>. Here each effector gene codes for a cell-type-specific protein required to generate the cell-specific output. These effector genes are all transcribed specifically in the given cell type in response to a small number of regulatory factors, which are themselves the output of the controlling specification GRN. Every effector gene of the battery is specifically controlled by these inputs. The immediate drivers of the battery shown cross-regulate (as is often the case). **b**, Structure that may be typical of morphogenetic effector gene cassettes. Here the output of the specification GRN is used to control transcription of only a minor fraction of key effector genes, and these in some way trigger or nucleate the process. But many of the proteins required for the function are widely expressed.

linkages to many genes encoding transcriptional activators and repressors<sup>59,60</sup>, including a polycomb repressor that in turn targets regulatory genes associated with various differentiation states<sup>61</sup>. But also among the immediate targets of the triple feedback loop is the *Rest* gene, which encodes a factor that directly represses neurogenic differentiation genes<sup>62</sup>. This circuit is the mirror image of gene battery activation circuits.

### Structure/function relations for GRNs controlling diverse kinds of biology

The downstream effector gene cassettes required for development include those executing morphogenetic cell biology functions, as well as differentiation gene batteries. A distinction is that by definition, differentiation genes are expressed cell type-specifically, whereas genes required for functions such as motility, ingression, invagination, cell division, convergent extension, tube formation, branching, shape remodelling, epithelial-mesenchyme transition, etc., may be deployed in many diverse cell types and many diverse contexts in development.

If we imagine a canonical differentiation gene battery to be structured as in Fig. 2a, how different will be the topology of a morphogenetic gene cassette? One possible clue comes from various studies on GRN linkages that execute transcriptional control of cell replication in developing systems. The spatial patterns of cell replication of course affect morphology, because the size and shape of given portions of a structure depend on the number of rounds of cell division mediated by the regulatory state in each developing region. In several cases the exact outputs of a developmental GRN that specifically control cell cycle activity have been determined. For

example in developing pituitary, several linkages from the specification GRN directly control proliferation<sup>63</sup>: the *Pitx1* gene provides inputs into the cyclin D1 gene; the *Six1* gene acts to repress expression of a cell cycle arrest kinase; and *Six1* plus other factors of the pituitary regulatory state activate *c-myc* (also known as *Myc*). In the developing zebrafish eye the GRN linkage to cell cycle control is regulation of *cyclin D1* and *c-myc* (also known as *myca/mycb*) by the *meis1* regulatory gene<sup>64</sup>. Thus, so to speak, these GRNs deploy the complex process of cell division by pressing a small number of regulatory 'buttons'.

Perhaps only a subfraction of the effector genes in a morphogenetic gene cassette are transcriptionally regulated by direct inputs from the upstream GRN. This concept emerged from a study of the migration of heart precursor cells in developing *Ciona*<sup>65</sup>, one of the few system-level investigations we have into the transcriptional control of a morphogenetic function. A large number of cell biology genes participate in the processes of membrane protrusion and motility required for heart cell migration, but most of these genes are widely expressed. Migratory activity is specifically deployed by transcriptional activation of the *rhoDF* gene, which encodes a key required GTPase, and it is this gene which is directly controlled by the *cis*-regulatory outputs of the upstream GRN. The same principle is evident in a study of trichome formation in *Drosophila*<sup>66</sup>. Here again, an extensive patterning GRN lies upstream, and determines the location of the morphological features and its cellular progenitors. The remodelling of epidermal cell shape to produce trichomes (or alternately, smooth cuticle) is controlled by expression of the regulatory gene *shavenbaby* (also known as *ovo*), and some of its direct effector gene targets are known. But these are again only a fraction of the total genes whose products are required to build the trichome. If these examples are a guide, the wiring of differentiation gene batteries, in which every downstream gene is a specific target of the GRN (Fig. 2a), is distinct from the way morphogenetic gene cassettes may be wired (Fig. 2b). Many of the genes contributing to a morphogenetic cell biology process may be widely expressed and only a few key 'button' genes that functionally nucleate the whole process are transcriptionally controlled by GRN outputs, to deploy the process spatially. Were this a general result, it would promise the existence of simple regulatory levers by which morphogenetic cassettes could be re-deployed, either in evolution or in re-engineering projects, to which we return below.

A uniquely explanatory GRN analysis of innate immunity response mechanisms in dendritic cells, following stimulation of Toll-like receptors (TLRs)<sup>67</sup>, shows how a classic physiological response is programmed at the genomic level. Stimulation of TLRs 2, 3 and 4 with various agonists activates two partly overlapping response programs of effector gene expression, in other words an antiviral program and an inflammatory program. This study included all regulatory genes specifically involved in the process, and the architecture of the GRN was based on a comprehensive, quantitative perturbation analysis, using small hairpin RNAs (shRNAs) to block regulatory gene transcription, although no direct *cis*-regulatory validation of the GRN structure was reported. Several interesting differences and also similarities emerge in comparing the structure of this physiological response GRN to that of the developmental GRNs considered above. A salient similarity is in the structure of the effector gene sets. Like many differentiation genes, the TLR response effector gene sets are largely wired to their drivers in coherent feed forward loops. Another now familiar feature is the use of positive feedback that will lock down the regulatory state following a transient input, here between *stat* genes high up in the antiviral response GRN hierarchy. This GRN is of moderate depth: downstream of the *stat* genes are three other regulatory genes linked to the *stat* genes and to one another by cross-regulatory interactions, and downstream of these in turn are further regulatory genes, and then the effector genes. A further device in these GRNs that also is often used in development, is exclusion of the alternate regulatory state by specific cross-repression, once one of the pathways is active. The depth of the inflammatory hierarchy is only that of the feed forward circuitry. Physiological systems are homeostatic, and a special feature of this one is a self-cancelling repression circuit the



sequence-specific basis of which is, however, yet unknown. Some years ago a prescient analysis predicted that in general, developmental GRNs which control progressive irreversible regulatory processes would have considerably greater depth than does reversible physiological response circuitry<sup>68</sup>, and this turns out to be exactly true.

One way of summarizing the result of a comparative meta-analysis of GRNs controlling diverse kinds of biological processes is to consider their similarities and differences in the same terms: they are similar in that all the GRNs considered here are modular constructs of a basic repertoire of sub-circuit topologies (Table 1); but they differ in their global hierarchical organization, which reasonably reflects the biological jobs they execute.

### Insights into process from mathematical models of GRNs and sub-circuits

Space confines the following discussion to recently conceived models based *ab initio* on experimentally generated, largely validated network topologies. The major focus is on how, or whether, mathematical analyses of the models has succeeded in enriching our understanding of the biological functionalities of the observed circuitry.

Beginning with a known network topology, the common objective is to generate a dynamic mathematical model, either using continuous (ordinary differential equations or ODE) or Boolean approaches<sup>69</sup>. For large scale temporal models of embryonic spatial specification systems involving many genes and interactions, this often involves a great number of unmeasured parameters, and epistemological issues immediately arise. In many such works arbitrary parameter values are systematically explored until the expected results emerge, but this is inherently at least a partially circular logic, since it assumes *a priori* that the model is right. Of course where there are applicable experimental measurements of the output kinetics, the model is better constrained, but then the novelty of the biological insights that can be expected is limited because both the input relationships and the results are assumed. *Drosophila* gap gene expression in the syncytial embryo provides the best known large developmental data set thus far subjected to mathematical kinetic analysis<sup>70–72</sup>. Extensive genetic and *cis*-regulatory data partially specify the embryonic interaction networks of these genes<sup>73–75</sup>. Mathematical models were built assuming the network topologies proposed in prior work<sup>70,73</sup>, and fit to a very high quality set of quantitative kinetic measurements which capture the empirical dynamics of changing gap expression patterns in the pre-cellularization 13th–14th cleavage cycle<sup>71,72</sup>. There were two outcomes relevant to the structure/function relationships of this developmental GRN: First, a dynamic image of how the gap gene transcription factors operate emerged, illuminating what might be called the cell biology of the process (were there cells). Second, the analysis suggested several additions and corrections of unresolved details of gap gene interactions. But largely the outcome was just that if one does the math and the measurements, everything turns out to make sense.

An important area of developmental biology in which modelling has contributed novel mechanistic understanding is transcriptional response to signals. We cannot here deal with the many studies focused on dynamic spatial distributions of signal ligands *per se*. But mathematical analyses of models capturing transcriptional network circuitry downstream of intercellular signalling have illuminated developmental signal response in multiple ways. These models concern smaller and well constrained sub-circuits, rather than whole GRNs, and often either parameters can be reasonably approximated, or dimensionless approaches can be found. The signal-driven transcriptional patterning process by which the two dorsal respiratory appendages on the roof of the *Drosophila* egg are positioned affords an example<sup>76</sup>. An experimentally based network circuitry animated by spatially confined epidermal growth factor (EGF) and Dpp signalling was used to produce a dynamic mathematical model which satisfactorily interprets the changing pattern of expression of a key gene of the pro-appendage regulatory state in dorso-anterior follicle cells. The model thus explains how this system generates and positions the bilateral spots of gene expression where the

appendages will form, which is not otherwise transparent. Furthermore, in consequence of a conflict between prediction and experimental observations, the analysis required a hitherto unsuspected positive feedback loop by which Dpp controls expression of its own receptor. A second example concerns transcriptional interpretation of graded hedgehog (Hh) signals in the developing neural tube, which results in a ventral to dorsal series of spatial regulatory state domains each of which gives rise to certain neuronal types<sup>34</sup>. When experimental measurements of signal intensity over time in the various transcriptional domains were analysed mathematically<sup>77,78</sup>, it emerged that the successive ventral to dorsal transcriptional domains are defined by the integrals over duration and intensity of Hh signalling, rather than simply on 'morphogen concentration', as always assumed previously. A third example<sup>79,80</sup> relates to the Wnt signalling required in *Xenopus* embryos to activate key regulatory genes of the dorsal organizer. Experimental perturbations of this canonical developmental signalling system showed that this system responds to the ratio of the (signal) input at some given time, to its level when the signalling began ('fold change'), and not to absolute signal level (the same phenomenon is often seen in other contexts, for example, sensory physiology). A predicted explanation in terms of network sub-circuit topology was then derived from a dynamic mathematical analysis of the incoherent feed forward sub-circuit<sup>80</sup>, which showed that this commonplace sub-circuit possesses the capacity to respond to fold change in input magnitude, rather than to absolute input magnitude.

As noted above, another general area in which modelling has illuminated process in respect to given sub-circuit topology is in binary cell fate choice, following a precursor phase in which both regulatory states are weakly expressed. Here the repeatedly observed sub-circuit structure features the opposition of two antagonistic repressors, each, if highly expressed, capable of shutting off the alternative regulatory state and generating its own, and each animated by inputs that reflect the external need for its pathway. A canonical approach to dynamic mathematical modelling of this type of sub-circuit has been repeatedly applied, based essentially on treating transcriptional activation abstractly as a catalytic Michaelis–Menton process, and repression in the same vein (for example, refs. 38, 81). The object is to demonstrate that these 'duelling repressor' sub-circuit topologies indeed encode regulatory systems that are capable of hysteretically moving from the precursor state to one or the other terminal regulatory states, depending on the inputs the system receives. But a problem with this approach is that as conventionally formulated, the bi-stable mathematical behaviour requires the completely ad hoc assumption of large exponential (Hill) coefficients in the repression functions (that is, coefficients  $>2$ , and often much larger values have to be assumed in order to generate the expected behaviour). Although Hill coefficients of these magnitudes physically imply cooperativity, or additional (unknown) reactions, they are customarily inserted in the computations despite lack of any direct biological evidence for cooperativity or other physical features that would justify them. Indeed, in one recent study of another very similarly wired haematopoietic choice system<sup>82</sup>, the erythroid/myeloid fate choice, it was pointed out that the specific, mutually repressive *cis*-regulatory interactions which were obtained are known not to be multimeric and cooperative, nor is there any other biochemical justification for high Hill coefficients. Instead an alternative regulatory architecture was considered, the dynamic mathematical analysis of which resulted in the prediction that the sub-circuit should include in addition to the antagonistic repressors another gene or genes operating according to specified network linkages. The latter work<sup>82</sup>, furthermore, used a now classic probabilistic thermodynamic treatment of *cis*-regulatory transcription factor binding<sup>83</sup> that is directly based on transcription factor–DNA interaction physical chemistry. This same thermodynamic approach to modelling *cis*-regulatory transcription factor binding has been used for analysis of an entirely different type of sub-circuit operating at the initial developmental appearance of pluripotential haematopoietic stem cells<sup>28,47</sup>. This sub-circuit consists of three positively active genes. There are no cross-regulating repressors in this sub-circuit, and the three genes are

linked by multiple positive auto- and cross-regulatory linkages. In life and in the model, extrinsic signals switch it irreversibly into an active state; otherwise, if one node is inhibited, it remains off<sup>47</sup>. Thus there are multiple different designs that confer signal-dependent bi-stability.

The thermodynamic binding approach<sup>83</sup> was also used earlier for dynamic modelling of sea urchin embryo gene cascades<sup>84</sup>. The important insight emerged that in a cascade where a given gene activates a second downstream gene, significant expression of the second gene occurs long before the product of the first gene reaches steady state, and the whole dynamic system operates in a 'forward drive mode' relatively insensitive to levels of upstream activators. The kinetics of such embryonic regulatory interactions are not narrowly determinate, as emphasized by the kinetic 'sloppiness' of a process which operates successfully at different rates at different temperatures within and between similar species, and in which there is a significant range in the concentrations of many transcription factors embryo to embryo<sup>85</sup>.

For other situations in embryonic development where the object is to encompass a complex, large scale spatial specification system rather than to follow a given small domain or cell type through time, conventional, stand-alone dynamical analysis is the wrong tool for the job that really needs to be done. Returning to Table 1, for example, we see that there are several kinds of spatial specification sub-circuit, that in cellular early embryos produce novel spatial regulatory state patterns, for example X,1 – X spatial processing sub-circuits and AND spatial logic processors. These, and indeed many other embryonic specification processes that define multicellular territorial regulatory states, result in a progressive Boolean-like pattern of diverse regulatory states confronting one another sharply across territorial cellular boundaries. A model that would capture what the GRN really does must address this kind of outcome, capturing the encoded input information-processing behaviour at each *cis*-regulatory module of the GRN.

Current developmental GRNs mainly concern, on the one hand, far upstream hierarchical transactions that essentially execute regulatory state pattern formation, or on the other, far downstream differentiation gene batteries and their immediate governance. These will have to be much better linked, so that we have a continuous understanding of the control systems from the top of the hierarchy to all the effector genes of a developing system. This kind of global GRN will be much larger than anything we have at present. Other kinds of global GRNs are on the horizon as well, such as those that encompass all parts of a developing embryo through time. Experimentally validated GRNs that include complete large regulatory systems will present enormous computational challenges for modelling, presentation, logic analysis and modular abstraction.

## Developmental GRNs and evolutionary mechanism

Because development of the body plan is caused by the operation of GRNs, evolutionary change in the body plan is change in GRN structure occurring over deep time<sup>86–88</sup>. Evolution and development emerge as twin outputs of the same mechanistic domain of regulatory system genomics. It is therefore to be expected that, at the level of GRN structure, each would illuminate the other, and so indeed they already do in several concrete ways. To start, it is obvious that if there is indeed a finite repertoire of network sub-circuits used to effect development, the evolution of development has to be considered as the process of assembly, reassembly, and redeployment of these sub-circuits. This general idea will become directly testable by widespread evolutionary comparisons, as the GRNs underlying the development of diverse animal forms become available. Structural comparison of GRNs between forms of known phylogenetic relation in turn reveals the modularity of GRN structure, by revealing sub-circuit boundaries, as when a sub-circuit is inserted wholesale into a new GRN context<sup>89</sup>. Furthermore, the sub-circuits of which GRNs are composed change during evolution at different rates, highlighting the linkages belonging to the most conserved sub-circuit in a GRN comparison. As discussed elsewhere, in general the

oldest GRN features are certain differentiation gene batteries<sup>3,88</sup>, which are eumetazoan (cnidarian + bilaterian) in distribution. In contrast, the morphogenetic programs that pattern each form of body plan are by definition clade-specific<sup>88</sup>. Certain remarkably conserved regulatory sub-circuits that are located near the top of developmental GRN hierarchy may serve to lock down developmental process specific to given phyla or classes (GRN kernels)<sup>86–91</sup>. Thus GRNs are historically as well as structurally and functionally modular, in that they are a mosaic of sub-circuits of diverse antiquity and phylogenetic distribution. Systematic exploration of phylogenetically related GRNs at different distances is valuable not only to discover the evolutionary origins of each sub-circuit, but also to reveal which kinds of sub-circuits and linkages are inherently flexible and which not. This brings us to the most important point for the future. In order to probe control of spatial regulatory state, laboratory strategies will need to be designed for changing GRNs by insertion of network regulatory apparatus into developing systems. But this is the same kind of change that happened in evolution, and the results will be mutually informative. Thus a practical convergence is on the horizon. Re-engineering spatial developmental processes, and recreating evolutionary processes, while different in motivation, will both depend on fundamental understanding and experimental manipulation of the structure/function relations of developmental GRNs.

The processes we have been discussing, development and evolution of the body plan, and execution of physiological responses, devolve causally from the regulatory genome. We need to understand GRNs because they encompass the primary output of the regulatory genome, itself the fundamental and unique outcome of more than 600 million years of animal evolution<sup>88</sup>.

1. Oliveri, P., Tu, Q. & Davidson, E. H. Global regulatory logic for specification of an embryonic cell lineage. *Proc. Natl Acad. Sci. USA* **105**, 5955–5962 (2008).  
**This paper provides proof of principle that if a developmental GRN is essentially complete, then it provides causal explanations for the biological functions of the process it controls.**
2. Peter, I. S. & Davidson, E. H. Modularity and design principles in the sea urchin embryo gene regulatory network. *FEBS Lett.* **583**, 3948–3958 (2009).  
**This paper presents the latest comprehensive review of the sea urchin endomesoderm GRN, so far the most extensively validated large scale embryonic GRN, with special emphasis on the topologies of its spatial control sub-circuits.**
3. Davidson, E. H. *The Regulatory Genome. Gene Regulatory Networks in Development and Evolution* (Academic Press/Elsevier, 2006).
4. Alon, U. Network motifs: theory and experimental approaches. *Nature Rev. Genet.* **8**, 450–461 (2007).
5. Mangan, S. & Alon, U. Structure and function of the feed-forward loop network motif. *Proc. Natl Acad. Sci. USA* **100**, 11980–11985 (2003).
6. Davidson, E. H. Network design principles from the sea urchin embryo. *Curr. Opin. Genet. Dev.* **19**, 535–540 (2009).
7. Ma, W., Trusina, A., El-Samad, H., Lim, W. A. & Tang, C. Defining network topologies that can achieve biochemical adaptation. *Cell* **138**, 760–773 (2009).
8. Peter, I. S. & Davidson, E. H. The endoderm gene regulatory network in sea urchin embryos up to mid-blastula stage. *Dev. Biol.* **340**, 188–199 (2010).
9. Oliveri, P. & Davidson, E. H. Built to run, not fail. *Science* **315**, 1510–1511 (2007).
10. Koide, T., Hayata, T. & Cho, K. W. Y. *Xenopus* as a model system to study transcriptional regulatory networks. *Proc. Natl Acad. Sci. USA* **102**, 4943–4948 (2005).
11. Maduro, M. F. Structure and evolution of the *C. elegans* embryonic endomesoderm network. *Biochim. Biophys. Acta* **1789**, 250–260 (2009).
12. Chan, T.-M. *et al.* Developmental gene regulatory networks in the zebrafish embryo. *Biochim. Biophys. Acta* **1789**, 279–298 (2009).
13. Morley, R. H. *et al.* A gene regulatory network directed by zebrafish No tail accounts for its roles in mesoderm formation. *Proc. Natl Acad. Sci. USA* **106**, 3829–3834 (2009).
14. Stathopoulos, A. & Levine, M. Genomic regulatory networks and animal development. *Dev. Cell* **9**, 449–462 (2005).
15. Hong, J.-W., Hendrix, D. A., Papatsenko, D. & Levine, M. S. How the Dorsal gradient works: insights from postgenome technologies. *Proc. Natl Acad. Sci. USA* **105**, 20072–20076 (2008).  
**This review summarizes work regulatory control of Dorsal target genes expressed spatially along the dorsal/ventral axis of the syncytial *Drosophila* embryo.**
16. Liberman, L. M. & Stathopoulos, A. Design flexibility in *cis*-regulatory control of gene expression: synthetic and comparative evidence. *Dev. Biol.* **327**, 578–589 (2009).  
**This paper presents a novel experimental evidence of cis-regulatory design features in the syncytial dorsal-ventral *Drosophila* specification system.**



17. Levine, M. & Davidson, E. H. Gene regulatory networks for development. *Proc. Natl Acad. Sci. USA* **102**, 4936–4942 (2005).
18. Ochoa-Espinosa, A., Yu, D., Tsigiris, A., Struffi, P. & Small, S. Anterior-posterior positional information in the absence of a strong Bicoid gradient. *Proc. Natl Acad. Sci. USA* **106**, 3823–3828 (2009).
- This paper provides experimental evidence that the anterior/posterior specification system of the *Drosophila* embryo is controlled by a network of gene interactions rather than only by quantitative positional values of Bicoid.**
19. Liberman, L. M., Teves, G. T. & Stathopoulos, A. Quantitative imaging of the Dorsal nuclear gradient reveals limitations to threshold-dependent patterning in *Drosophila*. *Proc. Natl Acad. Sci. USA* **106**, 22317–22322 (2009).
20. Huang, A. M., Rusch, J. & Levine, M. An anteroposterior Dorsal gradient in the *Drosophila* embryo. *Genes Dev.* **11**, 1963–1973 (1997).
21. Saka, Y. & Smith, J. C. A mechanism for sharp transition of morphogen gradient interpretation in *Xenopus*. *BMC Dev. Biol.* **7**, 47–55 (2007).
22. Davidson, E. H. *Genomic Regulatory Systems: Development and Evolution* (Academic Press/Elsevier, 2001).
23. Su, Y.-H. *et al.* A perturbation model of the gene regulatory network for oral and aboral ectoderm specification in the sea urchin embryo. *Dev. Biol.* **329**, 410–421 (2009).
24. Nikitina, N., Sauka-Spengler, T. & Bronner-Fraser, M. Dissecting early regulatory relationships in the lamprey neural crest gene network. *Proc. Natl Acad. Sci. USA* **105**, 20083–20088 (2008).
25. Woodland, H. R. & Zorn, A. M. The core endodermal gene network of vertebrates: combining developmental precision with evolutionary flexibility. *Bioessays* **30**, 757–765 (2008).
26. Cvekl, A. & Duncan, M. K. Genetic and epigenetic mechanisms of gene regulation during lens development. *Prog. Retin. Eye Res.* **26**, 555–597 (2007).
27. Kumar, J. P. The molecular circuitry governing retinal determination. *Biochim. Biophys. Acta* **1789**, 306–314 (2009).
28. Pimanda, J. E. *et al.* Gata2, Fli1, and Scl form a recursively wired gene-regulatory circuit during early hematopoietic development. *Proc. Natl Acad. Sci. USA* **104**, 17692–17697 (2007).
29. Smith, P. A. & Mango, S. E. Role of T-box gene *tbx-2* for anterior foregut muscle development in *C. elegans*. *Dev. Biol.* **302**, 25–39 (2007).
30. Cripps, R. M. & Olson, E. N. Control of cardiac development by an evolutionarily conserved transcription network. *Dev. Biol.* **246**, 14–28 (2002).
31. Reim, I., Mohler, J. P. & Frasch, M. *Tbx20*-related genes, *mid* and *H15* are required for *tinman* expression, proper patterning, and normal differentiation of cardioblasts in *Drosophila*. *Mech. Dev.* **122**, 1056–1069 (2005).
32. Albert, R. & Othmer, H. G. The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *Drosophila melanogaster*. *J. Theor. Biol.* **223**, 1–18 (2003).
33. Nishi, Y., Ji, H., Wong, W. H., McMahon, A. P. & Vokes, S. A. Modeling the spatio-temporal network that drives patterning in the vertebrate central nervous system. *Biochim. Biophys. Acta* **1789**, 299–305 (2009).
34. Vokes, S. A., Ji, H., Wong, W. H. & McMahon, A. P. A genome-scale analysis of the cis-regulatory circuitry underlying sonic hedgehog-mediated patterning of the mammalian limb. *Genes Dev.* **22**, 2651–2663 (2008).
35. Ririe, T. O., Fernandes, J. S. & Sternberg, P. W. The *Caenorhabditis elegans* vulva: A post-embryonic gene regulatory network controlling organogenesis. *Proc. Natl Acad. Sci. USA* **105**, 20095–20099 (2008).
36. Graf, T. & Enver, T. Forcing cells to change lineages. *Nature* **462**, 587–594 (2009).
- This review comprehensively traverses the process of terminal lineage fate choice in pluripotential hematopoietic systems.**
37. Swiers, G., Patient, R. & Loose, M. Genetic regulatory networks programming hematopoietic stem cells and erythroid lineage specification. *Dev. Biol.* **294**, 525–540 (2006).
38. Laslo, P. *et al.* Multilineage transcription priming and determination of alternate hematopoietic cell fates. *Cell* **126**, 755–766 (2006).
- This paper exemplifies a commonly used mathematical approach invoking bi-stable state kinetics to explain lineage choice.**
39. Smith, J. & Davidson, E. H. Gene regulatory network subcircuit controlling a dynamic spatial pattern of signaling in the sea urchin embryo. *Proc. Natl Acad. Sci. USA* **105**, 20089–20094 (2008).
40. Zhang, P. *et al.* Negative cross-talk between hematopoietic regulators: GATA proteins repress PU.1. *Proc. Natl Acad. Sci. USA* **96**, 8705–8710 (1999).
41. Huang, S., Guo, Y.-P., May, G. & Enver, T. Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. *Dev. Biol.* **305**, 695–713 (2007).
42. Stopka, T., Amanatullah, D. F., Papetti, M. & Skoultschi, A. I. PU.1 inhibits the erythroid program by binding to GATA-1 on DNA and creating a repressive chromatin structure. *EMBO J.* **24**, 3712–3723 (2005).
43. Starck, J. *et al.* Functional cross-antagonism between transcription factors FLI-1 and EKL. *Mol. Cell. Biol.* **23**, 1390–1402 (2003).
44. Rothenberg, E. V. Decision by committee: new light on the CD4/CD8-lineage choice. *Immunol. Cell Biol.* **87**, 109–112 (2009).
45. Wang, L. & Bosselut, R. CD4–CD8 lineage differentiation: Thpok-ing into the nucleus. *J. Immunol.* **183**, 2903–2910 (2009).
46. Setoguchi, R. *et al.* Repression of the transcription factor Th-POK by Runx complexes in cytotoxic T cell development. *Science* **319**, 822–825 (2008).
47. Narula, J., Smith, A. M. & Gottgens, B. and Iqbal, O. A. Modeling reveals bistability and low-pass filtering in the network module determining blood stem cell fate. *PLoS Comp. Biol.* **6**, e1000771 (2010).
48. Hu, M. *et al.* Multilineage gene expression precedes commitment in the hematopoietic system. *Genes Dev.* **11**, 774–785 (1997).
49. Miyamoto, T. *et al.* Myeloid or lymphoid promiscuity as a critical step in hematopoietic lineage commitment. *Dev. Cell* **3**, 137–147 (2002).
50. Lagha, M. *et al.* *Pax3:Foxc2* reciprocal repression in the somite modulates muscular versus vascular cell fate choice in multipotent progenitors. *Dev. Cell* **17**, 892–899 (2009).
51. Johnson, R. J. Jr, Chang, S., Etchberger, J. F., Ortiz, C. O. & Hobert, O. MicroRNAs acting in a double-negative feedback loop to control a neuronal cell fate decision. *Proc. Natl Acad. Sci. USA* **102**, 12449–12454 (2005).
52. Vierbuchen, T. *et al.* Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **463**, 1035–1041 (2010).
53. Zhou, Q., Brown, J., Kanarek, A., Rajagopal, J. & Melton, D. A. *In vivo* reprogramming of adult pancreatic exocrine cells to  $\beta$ -cells. *Nature* **455**, 627–632 (2008).
54. Gilchrist, M. *et al.* Systems biology approaches identify ATF3 as a negative regulator of Toll-like receptor 4. *Nature* **441**, 173–178 (2006).
55. Hobert, O. Regulatory logic of neuronal diversity: Terminal selector genes and selector motifs. *Proc. Natl Acad. Sci. USA* **105**, 20067–20071 (2008).
56. Bröhl, D. *et al.* A transcriptional network coordinately determines transmitter and peptidergic fate in the dorsal spinal chord. *Dev. Biol.* **322**, 381–393 (2008).
57. Yun, K. & Wold, B. Skeletal muscle determination and differentiation: story of a core regulatory network and its context. *Curr. Opin. Cell Biol.* **8**, 877–889 (1996).
58. Pan, G. & Thomson, J. A. Nanog and transcriptional networks in embryonic stem cell pluripotency. *Cell Res.* **17**, 42–49 (2007).
59. Boyer, L. A. *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947–956 (2005).
60. Zhou, Q., Chipperfield, H., Melton, D. A. & Wong, W. H. A gene regulatory network in mouse embryonic stem cells. *Proc. Natl Acad. Sci. USA* **104**, 16438–16443 (2007).
61. Boyer, L. A. *et al.* Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**, 349–353 (2006).
62. Mortazavi, A., Chen Leeper Thompson, E., Garcia, S. T., Myers, R. M. & Wold, B. Comparative genomics modeling of the NRSF/REST repressor network: from single conserved sites to genome-wide repertoire. *Genome Res.* **16**, 1208–1221 (2006).
63. Zhu, X. & Rosenfeld, M. G. Transcriptional control of precursor proliferation in the early phases of pituitary development. *Curr. Opin. Genet. Dev.* **14**, 567–574 (2004).
64. Bessa, J. *et al.* *meis1* regulates *cyclin D1* and *c-myc* expression, and controls the proliferation of the multipotent cells in the early developing zebrafish eye. *Development* **135**, 799–803 (2008).
65. Christiaen, L. *et al.* The transcription/migration interface in heart precursors of *Ciona intestinalis*. *Science* **320**, 1349–1352 (2008).
- This paper presents direct evidence of the regulatory structure of a morphogenetic gene cassette, showing that only certain key genes are controlled by the upstream GRN while a majority are expressed anyway.**
66. Chanut-Delalande, H., Fernandes, I., Roch, F., Payre, F. & Plaza, S. Shavenbaby couples patterning to epidermal cell shape control. *PLoS Biol.* **4**, 1549–1561 (2006).
67. Amit, I. *et al.* Unbiased reconstruction of a mammalian transcriptional network mediating pathogen responses. *Science* **326**, 257–263 (2009).
- This paper presents the most complete analysis yet available of structure and function in a physiological GRN.**
68. Rosenfeld, N. & Alon, U. Response delays and the structure of transcription networks. *J. Mol. Biol.* **329**, 645–654 (2003).
69. Bolouri, H. *Computational Modeling of Gene Regulatory Networks – A Primer* (Imperial College Press, 2008).
70. Sánchez, L. & Thieffry, D. A logical analysis of the gap gene system. *J. Theor. Biol.* **211**, 115–141 (2001).
71. Jaeger, J. *et al.* Dynamical analysis of regulatory interactions in the gap gene system of *Drosophila melanogaster*. *Genetics* **167**, 1721–1737 (2004).
72. Perkins, T. J., Jaeger, J., Reintz, J. & Glass, L. Reverse engineering the gap gene network of *Drosophila melanogaster*. *PLoS Comp. Biol.* **2**, e051 (2006).
- This paper provides a comprehensive computational treatment of the *Drosophila* gap gene network using estimates of numerous constants obtained by high resolution imaging.**
73. Rivera-Pomar, R. & Jaekle, H. From gradients to stripes in *Drosophila* embryogenesis: filling in the gaps. *Trends Genet.* **12**, 478–483 (1996).
74. Kraut, R. & Levine, M. Mutually repressive interactions between the gap genes *giant* and *Krüppel* define middle body regions of the *Drosophila* embryo. *Development* **111**, 611–621 (1991).
75. Segal, E., Raveh-Sadka, T., Schroeder, M., Unnerstall, U. & Gaul, U. Predicting expression patterns from regulatory sequence in *Drosophila* segmentation. *Nature* **451**, 535–540 (2008).
76. Lembo, J., Yakoby, N. & Shvartsman, S. Y. Pattern formation by dynamically interacting network motifs. *Proc. Natl Acad. Sci. USA* **106**, 3213–3218 (2009).
77. Dessaud, E. *et al.* Neural dynamic assignment and maintenance of positional identity in the ventral neural tube by the morphogen Sonic hedgehog. *PLoS Biol.* **8**, e1000382 (2010).
- This paper provides a new insight into how positional values of Hedgehog ligand are used to set transcriptional thresholds.**
78. Ribes, V. & Briscoe, J. Establishing and interpreting graded Sonic Hedgehog signaling during vertebrate neural tube patterning: the role of negative feedback. *Cold Spring Harb. Perspect. Biol.* **1**, a002014 (2009).
79. Goentoro, L., Shoval, O., Kirschner, M. W. & Alon, U. The incoherent feedforward loop can provide fold-change detection in gene regulation. *Mol. Cell* **36**, 894–899 (2009).
- This analysis shows how a common GRN sub-circuit can operate to interpret relative changes in signal strength.**
80. Goentoro, L. & Kirschner, M. W. Evidence that fold-change, and not absolute level, of  $\beta$ -catenin dictates Wnt signaling. *Mol. Cell* **36**, 872–884 (2009).
81. Spooner, C. J. *et al.* A recurrent network involving the transcription factors PU.1 and Gfi1 orchestrates innate and adaptive immune cell fates. *Immunity* **31**, 576–586 (2009).

82. Chickarmane, V., Enver, T. & Peterson, C. Computational modeling of the hematopoietic erythroid-myeloid switch reveals insights into cooperativity, priming, and irreversibility. *PLoS Comp. Biol.* **5**, e1000268 (2009).  
**This paper presents an alternative computational treatment of lineage choice in a haematopoietic system.**
83. Shea, M. A. & Ackers, G. K. The  $O_R$  control system of bacteriophage lambda: A physical-chemical model for gene regulation. *J. Mol. Biol.* **181**, 211–230 (1985).
84. Bolouri, H. & Davidson, E. H. Transcriptional regulatory cascades in development: Initial rates, not steady state, determine network kinetics. *Proc. Natl Acad. Sci. USA* **100**, 9371–9376 (2003).  
**This paper models sea urchin regulatory cascade kinetics and demonstrates using measured constants that genes are successively activated long before any of the transcriptional functions attain steady state.**
85. Materna, S. C., Nam, J. & Davidson, E. H. High accuracy, high-resolution prevalence measurement for the majority of locally expressed regulatory genes in early sea urchin development. *Gene Expr. Patterns* **10**, 177–184 (2010).
86. Davidson, E. H. & Erwin, D. H. Gene regulatory networks and the evolution of animal body plans. *Science* **311**, 796–800 (2006).  
**This paper introduced the theory that highly conserved GRN sub-circuits account for the phylogenetic distribution of major characters of the animal body plan.**
87. Erwin, D. H. & Davidson, E. H. The evolution of hierarchical gene regulatory networks. *Nature Rev. Genet.* **10**, 141–148 (2009).
88. Davidson, E. H. & Erwin, D. H. An integrated view of Precambrian eumetazoan evolution. *Cold Spring Harb. Symp. Quant. Biol.* **74**, 65–80 (2009).
89. Gao, F. & Davidson, E. H. Transfer of a large gene regulatory apparatus to a new developmental address in echinoid evolution. *Proc. Natl Acad. Sci. USA* **105**, 6091–6096 (2008).
90. Hinman, V. F. & Davidson, E. H. Evolutionary plasticity of developmental gene regulatory network architecture. *Proc. Natl Acad. Sci. USA* **104**, 19404–19409 (2007).
91. Hinman, V. F., Yankura, K. A. & McCauley, B. S. Evolution of gene regulatory network architectures: Examples of subcircuit conservation and plasticity between classes of echinoderms. *Biochim. Biophys. Acta* **1789**, 326–332 (2009).
92. Bolouri, H. & Davidson, E. H. The gene regulatory network basis of the “community effect,” and analysis of a sea urchin embryo example. *Dev. Biol.* **340**, 170–178 (2010).

**Acknowledgements** I am grateful for the reviews of the manuscript by E. V. Rothenberg and I. S. Peter. This work was supported by NIH grants HD-37105 and GM-61005 and by the Lucille P. Markey Charitable Trust.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The author declares no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence should be addressed to the author (Davidson@caltech.edu).



# Light-avoidance-mediating photoreceptors tile the *Drosophila* larval body wall

Yang Xiang<sup>1</sup>, Quan Yuan<sup>1</sup>, Nina Vogt<sup>2</sup>, Loren L. Looger<sup>3</sup>, Lily Yeh Jan<sup>1</sup> & Yuh Nung Jan<sup>1</sup>

**Photoreceptors for visual perception, phototaxis or light avoidance are typically clustered in eyes or related structures such as the Bolwig organ of *Drosophila* larvae. Unexpectedly, we found that the class IV dendritic arborization neurons of *Drosophila melanogaster* larvae respond to ultraviolet, violet and blue light, and are major mediators of light avoidance, particularly at high intensities. These class IV dendritic arborization neurons, which are present in every body segment, have dendrites tiling the larval body wall nearly completely without redundancy. Dendritic illumination activates class IV dendritic arborization neurons. These novel photoreceptors use phototransduction machinery distinct from other photoreceptors in *Drosophila* and enable larvae to sense light exposure over their entire bodies and move out of danger.**

Light sensing is critical for animal life. Whereas image-forming visual perception allows animals to identify and track mates, predators and prey, non-image-forming functions regulate pupil reflex, phototaxis and circadian entrainment<sup>1,2</sup>. In addition to eyes<sup>1,2</sup>, extra-ocular photoreceptors exist<sup>1–5</sup>. For example, many eyeless or blinded animals can sense illumination of their body surfaces<sup>3–5</sup>. Birds possess deep-brain photoreceptors in their hypothalamus<sup>6</sup>, and extra-ocular photoreceptors are required for magnetic orientation of amphibians<sup>7</sup>. Recent studies demonstrate that eyeless animals such as *Caenorhabditis elegans* nonetheless have photoreceptors controlling light avoidance<sup>8–10</sup>.

*Drosophila* larvae spend most of the time feeding by digging into food. Light avoidance is a crucial behaviour to minimize body exposure. When tested in groups in a dark/light choice assay, *Drosophila* larvae prefer darkness<sup>11,12</sup>. This behaviour requires the pair of Bolwig organs on the larval head<sup>12</sup>; that is, primitive eye structures each comprised of 12 photoreceptors expressing Rh5 or Rh6, rhodopsins sensing blue and green light, respectively<sup>13</sup>.

## Cells besides Bolwig organs contribute to photoavoidance

We designed a photoavoidance assay for a single larva with sunlight-level intensities ( $\sim 1 \text{ mW mm}^{-2}$  in San Francisco on a clear day in June, consistent with previous reports<sup>14</sup>). Wild-type *Drosophila* larvae showed avoidance of a white light spot of  $0.57 \text{ mW mm}^{-2}$  (Fig. 1a and Supplementary Movie 1). Surprisingly, similar avoidance (Fig. 1b and Supplementary Movie 2) was exhibited by larvae with their Bolwig organs ablated by the pro-apoptotic gene *Head involution defective* (*Hid*; also called *Wrinkled* (*W*))<sup>15</sup> expressed via the Bolwig-organ-specific promoter *Glass Multimer Reporter* (*GMR*)<sup>16</sup> (Supplementary Fig. 1a). Lower light intensities elicited less photoavoidance of wild-type animals, and even less of Bolwig-organ-ablated animals (Fig. 1c). However, at light intensities of  $0.57 \text{ mW mm}^{-2}$  or higher, *GMR-Hid* larvae showed avoidance comparable to wild-type animals ( $P > 0.05$ ). Thus, although the Bolwig organs are responsible for dim light avoidance and dark congregation<sup>12</sup>, *Drosophila* larvae must contain extra-ocular photoreceptors.

Testing the wavelength dependence of photoavoidance using band-pass filters letting through ultraviolet ( $\sim 360 \text{ nm}$ ; Fig. 1d), violet

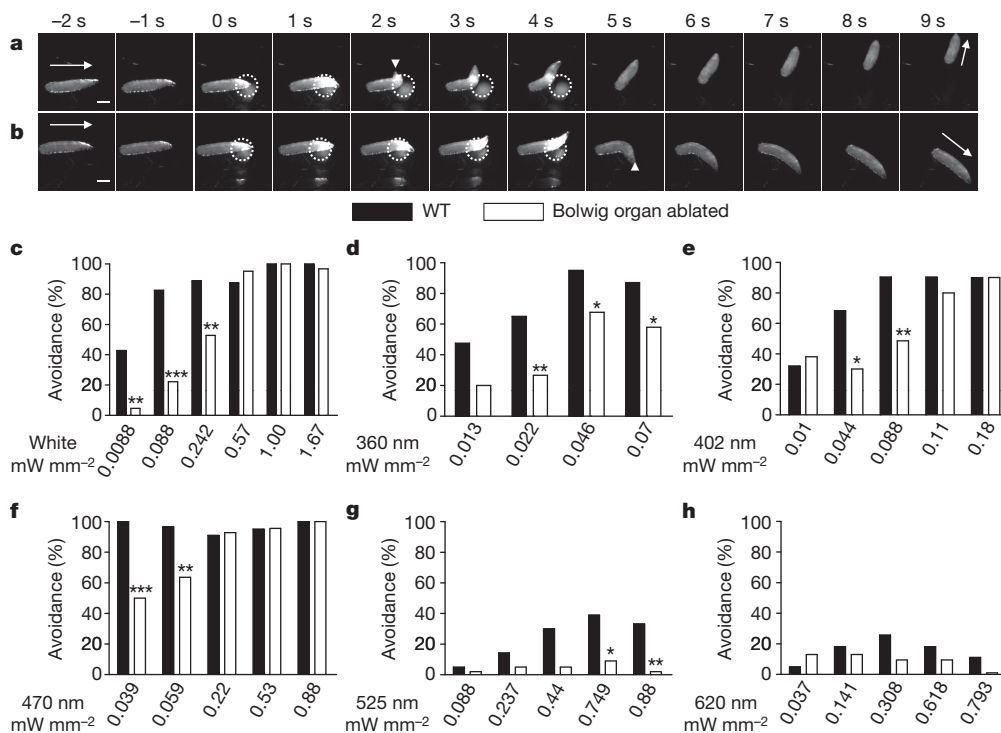
( $\sim 402 \text{ nm}$ ; Fig. 1e), blue ( $\sim 470 \text{ nm}$ ; Fig. 1f), green ( $\sim 525 \text{ nm}$ ; Fig. 1g) or red light ( $\sim 620 \text{ nm}$ ; Fig. 1h), we found that wild-type animals showed increased photoavoidance with higher light intensity (Fig. 1d–h), and were most sensitive to blue, violet and ultraviolet, and largely unresponsive to green and red light. Bolwig-organ-ablated animals showed less photoavoidance at low light intensity, but exhibited nearly normal avoidance response to high-intensity, short-wavelength light (Fig. 1d–f), demonstrating the existence of light-sensitive cells in addition to Bolwig organs. Because there was no detectable temperature increase associated with  $0.11 \text{ mW mm}^{-2}$  violet light (Supplementary Fig. 2)—which triggered avoidance in nearly 80% of the wild-type and *GMR-Hid* animals—and animals showed little response to high-intensity green or red light but strongly avoided low-intensity short-wavelength light (Fig. 1d–h), light avoidance probably involves wavelength-dependent photoreceptors but not local heating.

## Class IV neurons tiling larval body wall sense light

Given the report of diffusely distributed dermal photoreceptors triggering shadow reaction<sup>3–5</sup>, we tested whether sensory neurons in the larval body wall could be candidate photoreceptors. Using GCaMP3, a genetically encoded calcium indicator<sup>17–19</sup>, we found that blue light delivered for 5 s to the dorsal cluster (Fig. 2a, see Fig. 3c for whole larva image) generated a marked fluorescence increase specifically in the soma, axon and dendrites of ddaC, a class IV dendritic arborization neuron (Fig. 2b, e, f), but not in nearby sensory neurons (Fig. 2b, e, f). ddaC also responded to ultraviolet light (which also caused photobleaching), but not to green light (Fig. 2c–f). There were also  $\text{Ca}^{2+}$  increases specifically in class IV dendritic arborization neurons of the ventral and lateral cluster (*V'ada* and *VdaB*, respectively) in response to ultraviolet and blue light, but not green light (Supplementary Figs 3 and 4). Similar GCaMP3 fluorescence responses were seen in class IV dendritic arborization neurons in body segments from head to tail (Supplementary Fig. 5).

Extracellular recording further revealed a progressive increase in action potential frequency when the dorsal class IV dendritic arborization neuron ddaC was illuminated with increasing intensity of blue light (Fig. 2g), ultraviolet light and violet light, but not red light

<sup>1</sup>Howard Hughes Medical Institute, Departments of Physiology, Biochemistry, and Biophysics, University of California San Francisco, San Francisco, California 94158, USA. <sup>2</sup>Center for Developmental Genetics, New York University, New York, New York 10003, USA. <sup>3</sup>Howard Hughes Medical Institute, Janelia Farm Research Campus, Ashburn, Virginia 20147, USA.



**Figure 1 | Photoreceptors in addition to Bolwig organs contribute to photoavoidance.**

**a, b**, Examples of light avoidance of wild-type (**a**) and *GMR-Hid* (**b**) larvae exposed to white light ( $0.57 \text{ mW mm}^{-2}$ ) applied from 0 to 5 s. The light spot is indicated by the dotted circle. The arrow indicates the direction of larval locomotion; arrowheads at 2 s (**a**) and 5 s (**b**) indicate larval head turning. **c–h**, Percentage of animals avoiding white light (**c**), light of 360 nm (ultraviolet; **d**), 402 nm (violet; **e**), 470 nm (blue; **f**), 525 nm (green; **g**) and 620 nm (red; **h**) at different intensities. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , two-tailed Fisher exact test. Twenty to forty larvae were tested for each condition. Scale bar: 1 mm (**a, b**), shown at  $-2 \text{ s}$ .

(Supplementary Fig. 6). Responses were:  $340 \text{ nm} > 380 \text{ nm} > 402 \text{ nm} > 470 \text{ nm} \gg 525 \text{ nm}$  or  $620 \text{ nm}$  light (Fig. 2h).

The wavelength dependence of ddaC firing rate increase was similar to that observed with GCaMP3 imaging and the light avoidance behavioural assay. The latency between the onset of light stimulation and action potential burst firing decreased with higher light intensity, and was as short as 1 s with bright illumination (Supplementary Fig. 6). When illuminated with  $1.4 \text{ mW mm}^{-2}$  of white light (approximating sunlight), ddaC neurons in the dorsal cluster showed a significant firing increase (Fig. 2i). Similar robust activation of ventral (VdaB) and lateral (V'ada) class IV dendritic arborization neurons was induced by  $52.8 \text{ mW mm}^{-2}$  blue light (Supplementary Fig. 7a, b). The response of class IV dendritic arborization neurons was similar regardless of their location along the body axis (data not shown), as in the case of GCaMP3 imaging (Supplementary Fig. 5).

We did not observe any significant effects of light on firing rate of class I or III dendritic arborization neurons (Supplementary Fig. 7c, d,  $P > 0.05$ ). Because class I dendritic arborization neurons progressively increased their firing rate as the temperature was raised above  $30^\circ \text{C}$ , whereas class IV dendritic arborization neurons showed an abrupt increase of firing rate only above  $40^\circ \text{C}$  (Supplementary Fig. 8), thermal responses cannot account for the light-induced increase of firing in class IV but not class I dendritic arborization neurons. Moreover, application of  $10 \mu\text{M H}_2\text{O}_2$ , which elevates the reactive oxygen species (ROS) level in *Drosophila* larvae<sup>20</sup>, had no effect on the firing rate of class IV dendritic arborization neurons (Supplementary Fig. 9). These studies demonstrate that ultraviolet, violet and blue light activate class IV dendritic arborization neurons in an intensity-dependent manner. Responses occur at sunlight-level intensities, are not induced by heat or ROS, correlate with behaviour, and are confined to this specific class of sensory neurons throughout the animal.

### Light activates class IV neurons and dendrites in isolation

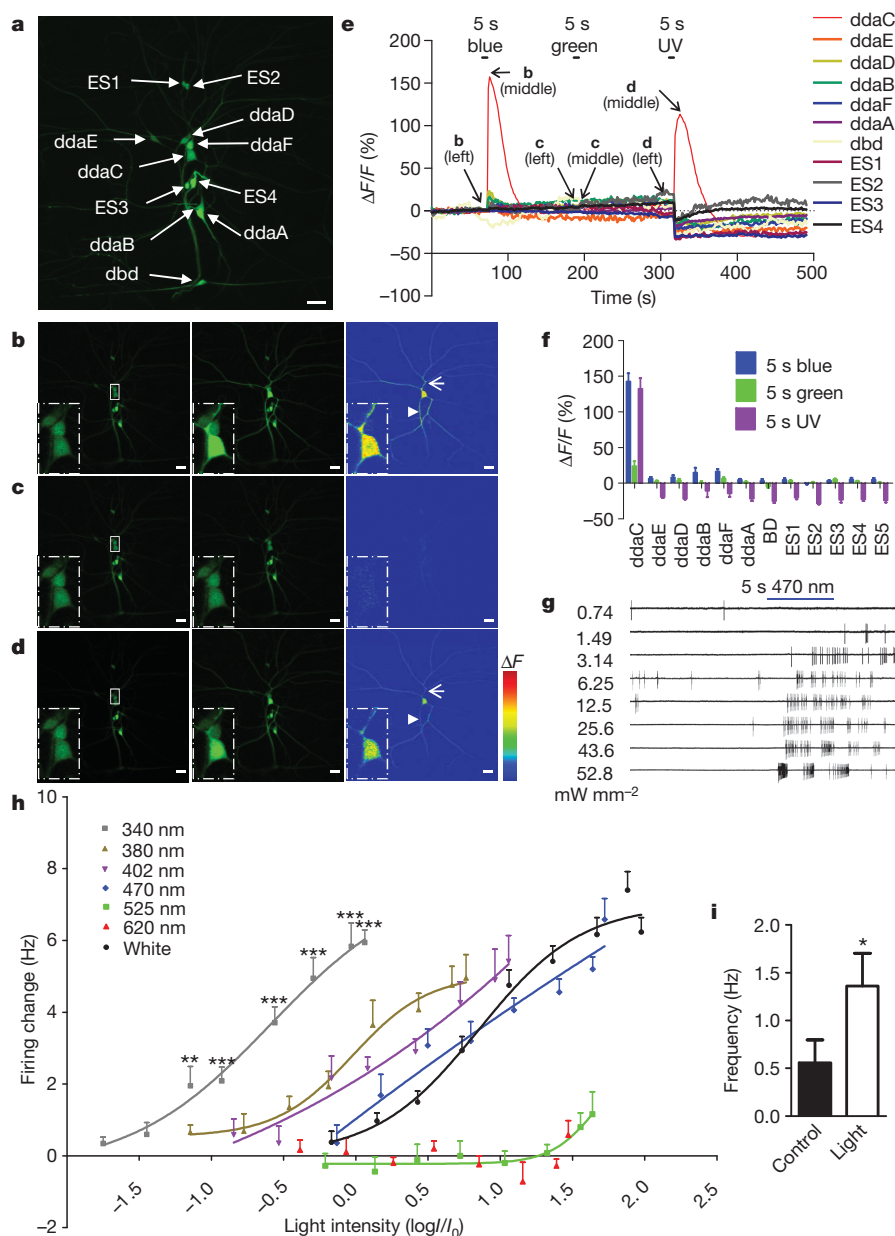
The dendritic arborization neurons have dendrites in contact with epithelial cells whereas their somas and axons are wrapped by glia<sup>21</sup>. To test whether class IV dendritic arborization neurons can sense light by themselves, we prepared primary neuronal cultures<sup>22,23</sup> from embryos expressing GCaMP3 and RFP specifically in class IV dendritic arborization neurons by means of *pickpocket-GAL4* (*ppk-GAL4*)<sup>24</sup>.

Ultraviolet and blue light illumination of isolated class IV dendritic arborization neurons generated a robust increase of GCaMP3 signals (Fig. 3a and Supplementary Fig. 10). In contrast, cultured class III dendritic arborization neurons expressing GCaMP3 and RFP via *19-12-GAL4* yielded no light response (Fig. 3b). Thus, class IV dendritic arborization neurons have the intrinsic ability to detect light.

Dendrites of class IV dendritic arborization neurons tile the larval body wall with non-overlapping but complete coverage of the dendritic field<sup>25,26</sup> (Fig. 3c). Illumination of only the dendrites of class IV dendritic arborization neurons (Fig. 3d) with ultraviolet, violet and blue light, but not green or red light, activated the neurons (Fig. 3e). The activation spectrum is similar to that for illumination of the entire class IV dendritic arborization neurons (Fig. 2h), indicating the presence of phototransduction machinery in the dendrites.

### Gr28b is critical for light transduction in class IV neurons

No defects in light response of class IV dendritic arborization neurons were found in available mutants of rhodopsins<sup>13,27</sup> and cryptochrome (*cry*)<sup>28</sup>, as well as a mutant in *no receptor potential A* (*norpA*), which encodes phospholipase C (PLC), downstream of rhodopsins<sup>29</sup> (Fig. 4a). We then tested the *Drosophila* homologue of *Lite-1*, a *C. elegans* light sensor<sup>8–10</sup>. The closest homologue of *lite-1* in *Drosophila* is gustatory receptor 28b (*Gr28b*), annotated as encoding a gustatory G-protein-coupled receptor. Several *Gr28b-GAL4* lines carrying different promoter regions revealed consistent expression in all class IV dendritic arborization neurons, two sensory neurons in the lateral body wall, plus several neurons in the ventral nerve cord (Supplementary Fig. 11), as reported previously<sup>30</sup>. To test for the functional role of *Gr28b* in the light-induced electrophysiological responses, we recorded from class IV dendritic arborization neurons in *Dmel*{Mi{ET1}Gr28b<sup>MB03888</sup> (*MiET1*) and *Dmel*{PBac{PB}Gr28b<sup>c01884</sup> (*PBac*)} larvae with P-element insertion into the *Gr28b* coding and intronic regions, respectively (<http://flybase.org/reports/FBgn0045495.html>). Whereas these P-element insertions did not alter the basal firing rate (data not shown), they caused a significant reduction in light-induced responses of class IV dendritic arborization neurons (Fig. 4b), as in hemizygous larvae carrying one *MiET1* allele and one deletion encompassing *Gr28b* (*Df*(2L)Exel7031; <http://flybase.org/reports/FBab0037910.html>) (Fig. 4c). The *MiET1* P-element inserts into the coding sequence



**Figure 2 | Light activates class IV dendritic arborization neurons.** **a**, Pre-stimulation image showing larval dorsal cluster sensory neurons (dbd, bipolar dendrite neuron; ddaD and ddaE, class I dendritic arborization neurons; ddaB, class II dendritic arborization neurons; ddaA and ddaF, class III dendritic arborization neurons; ddaC, class IV dendritic arborization neurons; ES, external sensory organ). Up is dorsal; right is anterior. For an atlas of the larval peripheral nervous system, see ref. 25. **b**, Responses of the dorsal cluster neurons in **a** to 5 s blue light (470 nm) illumination. The boxed area in the left panel and insets in all three panels show the somas of ddaC, ddaF and ddaD dendritic arborization neurons. Left, pre-stimulation; middle, post-stimulation; right, GCaMP3 intensity difference (middle panel minus left panel), with ddaC dendrites (arrow) and axon (arrowhead) marked. **c**, **d**, Similar experiments with 5 s green (546 nm; **c**) and ultraviolet light (365 nm; **d**) revealed ddaC activation by ultraviolet, but not green, light. Scale bar in **a–d**, 20  $\mu\text{m}$ ; colour scale in right panels of **b–d** shows dynamic range (0–4,095). **e**, Time course of somatic GCaMP3 signals of dorsal cluster neurons shown in **a–d**. Time frames are indicated. **f**, Summary of somatic fluorescence changes ( $\Delta F/F$ ) of dorsal cluster neurons in response to 5 s light stimulation,  $n = 7–16$ . **g**, Example firing traces of ddaC in response to 5 s 470 nm blue light. **h**, Summary of firing frequency changes (average frequency of 5 s before light exposure subtracted from average frequency during 5 s of light exposure) of ddaC induced by white, 340, 380, 402, 470, 525 and 620 nm light. For clarity, significance is only shown for the 340 nm curve. Light intensity is reported as the log of ( $I$  normalized to  $I_0 = 1 \text{ mW mm}^{-2}$ ). Green (525 nm) or red (620 nm) light has no effect ( $P > 0.05$ ).  $n = 5–9$ . **i**, Effect of  $1.4 \text{ mW mm}^{-2}$  white light on ddaC, average frequencies of 5 s before (control) and during the 5 s of light exposure (light) are plotted.  $n = 6$ . \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ ; two-tailed paired  $t$ -test. All error bars indicate s.e.m.

common to all reported transcripts, and its mobilization for excision restored the light-induced response in class IV dendritic arborization neurons (Fig. 4d). Moreover, knockdown of Gr28b expression with *UAS-RNAi* driven by *ppk-GAL4* caused an overall reduction of light response of class IV dendritic arborization neurons (Fig. 4e). Taken together, our data indicate that Gr28b is expressed in class IV dendritic arborization neurons, and is required for proper light responses. Whether Gr28b is the direct photosensing molecule awaits further experimentation.

Sequence analysis revealed that Gr28b has a rhodopsin-like structure plus one extra transmembrane segment (Supplementary Fig. 12), raising the question of whether the Gr28b-dependent light response involves G-protein signalling. To test whether G-protein signalling is required in class IV dendritic arborization neurons, we applied the myristoylated  $\beta\gamma$ -binding peptide mSIRK, and found that the light response in class IV dendritic arborization neurons was significantly reduced (Supplementary Fig. 13). Thus, G-protein signalling is probably involved in the light response of class IV dendritic arborization neurons, similar to findings in *C. elegans*<sup>10</sup>. We further tested cyclic nucleotide-gated (CNG) channels, which are known to act downstream of Lite-1 and G proteins in *C. elegans*<sup>10</sup>. Unlike in *C. elegans*, blocking CNG

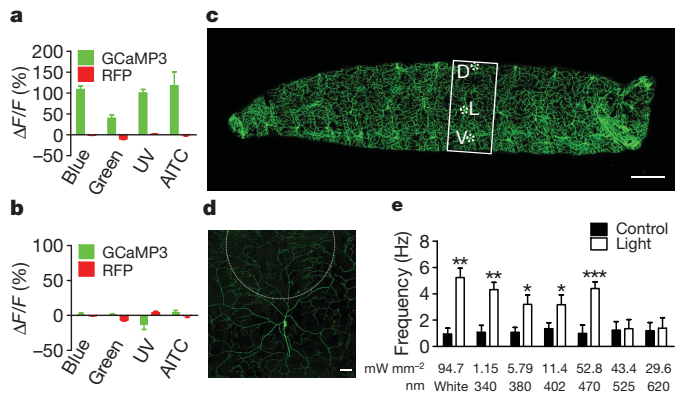
channels with *L-cis*-diltiazem in class IV dendritic arborization neurons had no effect on their light responses (Supplementary Fig. 14).

### TrpA1 is required in light transduction in class IV neurons

Transient receptor potential (TRP) channels were first identified and characterized in the *Drosophila* compound eye<sup>29,31</sup>, with TRP and TRP-like (*trpl*) having key roles in phototransduction<sup>29</sup>. However, our electrophysiological studies revealed no defects in the light response of class IV dendritic arborization neurons in *trpl* or *painless* mutant larvae (Supplementary Fig. 15).

TrpA1, a *Drosophila* homologue of mammalian TrpA, may function as a thermosensor in larvae and adults<sup>32–34</sup>, and a receptor for reactive electrophiles such as allyl isothiocyanate (AITC)<sup>35</sup>. A *TrpA1* mutant exhibited normal basal firing in class IV dendritic arborization neurons (data not shown), but no light-induced firing increase (Fig. 4f). As reported previously<sup>32</sup>, we detected strong TrpA1 immunoreactivity in several neurons in the larval brain but not in peripheral neurons (data not shown). We then performed MARCM (mosaic analysis with a repressible cell marker)<sup>36</sup>, and found no light response in class IV dendritic arborization neurons lacking TrpA1 (Fig. 4g), and a significant reduction of light-induced firing of the heterozygous class





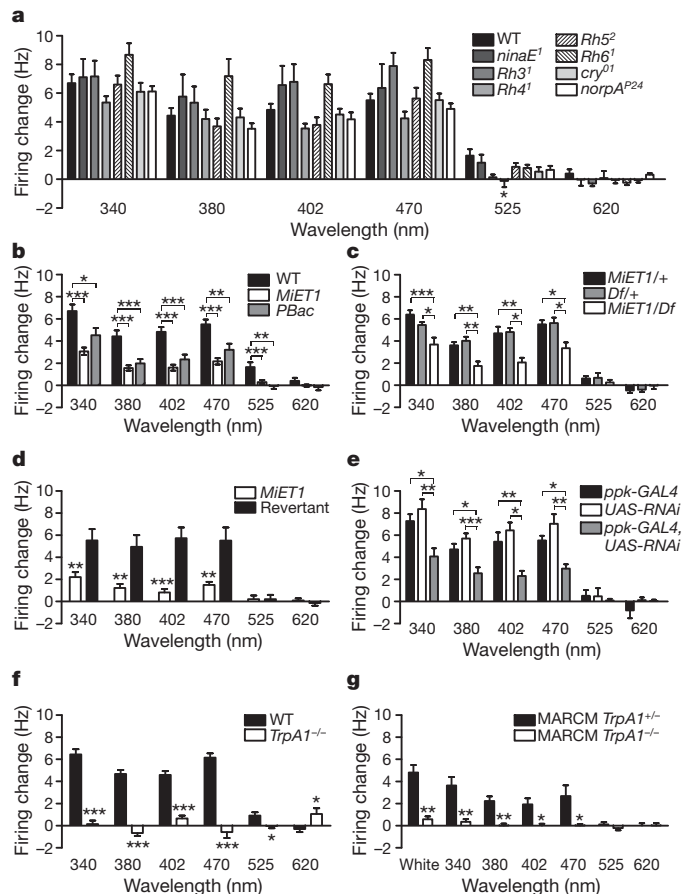
**Figure 3 | Cell-autonomous activation of class IV dendritic arborization neurons by light.** **a, b**, Quantification of somatic fluorescence changes ( $\Delta F/F$ ) in response to 5 s light and 100  $\mu$ M allyl isothiocyanate (AITC) stimulation of cultured class IV (**a**) and III (**b**) dendritic arborization neurons; RFP signals serve as control.  $n = 10$ –13 (light) and  $n = 4$  (AITC) in **a**,  $n = 9$  in **b**. **c**, Larva with class IV dendritic arborization neurons labelled with GFP by *ppk-GAL4*. Dendrites tile the body wall. Boxed area shows an abdominal hemi-segment; three dotted circles mark soma positions of D (dorsal, ddaC), L (lateral, V'ada) and V (ventral, VdaB) class IV dendritic arborization neurons, respectively. Up, dorsal; left, anterior. Scale bar, 200  $\mu$ m. **d**, Illumination of dendrites within the dotted circle of GFP-labelled ddaC dendrites. Up, dorsal. Scale bar, 50  $\mu$ m. **e**, Responses of ddaC with dendritic illumination.  $n = 5$ . \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ ; two-tailed paired *t*-test. All error bars indicate s.e.m.

IV dendritic arborization neurons (Supplementary Fig. 16), indicating that TrpA1 is present in levels below immunodetection, but nonetheless of functional importance. In support of this notion, AITC caused strong activation of class IV dendritic arborization neurons, and this activation was abolished in the *TrpA1* mutant (Supplementary Fig. 17a). Moreover, *TrpA1* RNAi expression specifically in class IV dendritic arborization neurons eliminated the light-induced firing change (Supplementary Fig. 18). Taken together, our observations suggest that TrpA1 is required cell-autonomously for light transduction in class IV dendritic arborization neurons.

Given the lack of AITC activation of class I or class III dendritic arborization neurons (Supplementary Fig. 17b), we expressed TrpA1 in class I dendritic arborization neurons and found that it conferred AITC sensitivity but not light response (Supplementary Fig. 17c, d), indicating that TrpA1 is not sufficient for light sensing. Because *trans*-heterozygotes carrying one mutant allele of *TrpA1* and one copy of the *MiET1* P-element insertion in the *Gr28b* gene showed reduced light response (Supplementary Fig. 19), it is likely that Gr28b and TrpA1 function in the same phototransduction pathway.

### Class IV neurons mediate light avoidance behaviour

To test whether class IV dendritic arborization neurons are involved in light avoidance, we genetically ablated class IV dendritic arborization neurons of third instar larvae by expressing the pro-apoptotic genes *Hid* (ref. 15) and *reaper* (*rpr*) (ref. 37) via *ppk-GAL4* (*ppk-GAL4; UAS-Hid,rpr*) (Supplementary Fig. 1). We also constructed a line lacking Bolwig organs as well as class IV dendritic arborization neurons (*UAS-Hid,rpr; GMR-Hid; ppk-GAL4*). Notably, both lines showed markedly decreased white-light-avoidance behaviour compared to wild-type and *GMR-Hid* (Bolwig-organ-ablated) larvae (Fig. 5a, b). Class IV dendritic-arborization-neurons-ablated animals showed a significant decrease of avoidance versus wild type, for all white light intensities tested (Fig. 5c–g). Ablation of class IV dendritic arborization neurons in animals lacking Bolwig organs produced a further decrease in white light avoidance (Fig. 5d–g). Avoidance of high-intensity ( $>0.57$  mW mm $^{-2}$ ) white light was normal when Bolwig organs were ablated in wild type (Fig. 5e–g), and in control strains with either *GAL4* or *UAS* (Fig. 5f). Taken together with similar

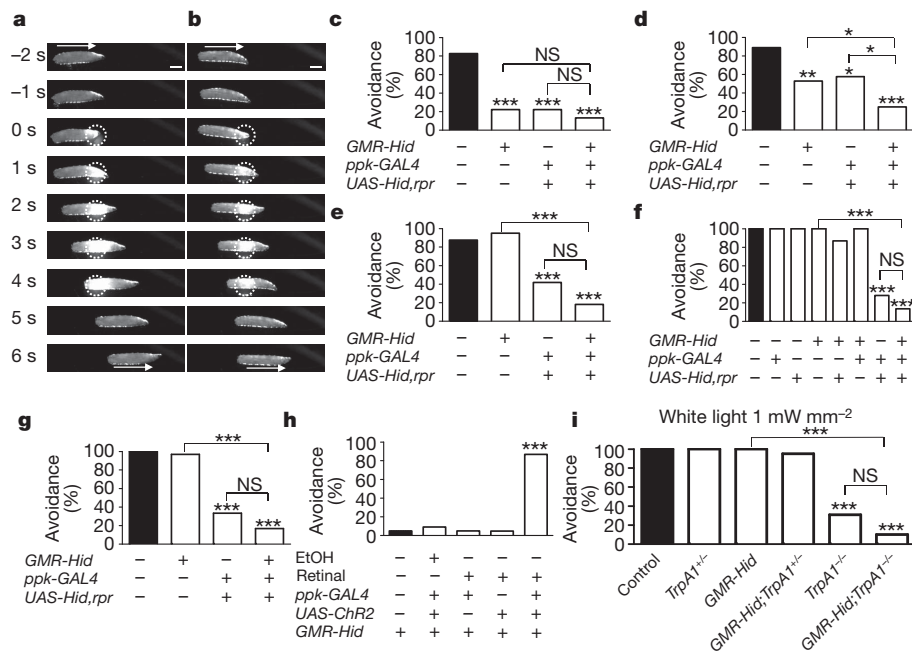


**Figure 4 | Gr28b and TrpA1 are essential for class IV dendritic arborization neuron light responses.** **a**, No significant defects were detected between wild-type and mutants of known phototransduction molecules with 340, 380, 402, 470, or 620 nm light.  $n = 5$ –10. **b**, Reduced light response of class IV dendritic arborization neurons in *MiET1* and *PBac* larvae.  $n = 8$ –29. **c**, Reduced light response of class IV dendritic arborization neurons in *MiET1*/deficiency larvae.  $n = 5$ –12. **d**, Precise excision of *MiET1* P-element insertion restores light response in class IV dendritic arborization neurons.  $n = 6$ –9. **e**, Reduced light responses of class IV dendritic arborization neurons with *Gr28b* RNAi knockdown.  $n = 5$ –8. **f**, Abolished light responses of class IV dendritic arborization neurons in *TrpA1* $^{-/-}$  mutants.  $n = 8$ –13. **g**, MARCM analysis of *TrpA1* $^{+/-}$  and *TrpA1* $^{-/-}$  class IV dendritic arborization neurons' response to light.  $n = 5$ –8. For **a**–**g**, Light intensities (mW mm $^{-2}$ ) are: 1.15 (340 nm), 5.79 (380 nm), 11.4 (402 nm), 52.8 (470 nm), 43.4 (525 nm), 29.6 (620 nm) and 94.7 (white). For **a**, **b**, **c**, **e**, \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ ; one-way ANOVA followed by a Bonferroni post test; for **d**, **f**, **g**, \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ ; two-tailed unpaired *t*-test. All error bars indicate s.e.m.

findings with ultraviolet, violet and blue light (Supplementary Figs 20–22), these results demonstrate that class IV dendritic arborization neurons are necessary to elicit photoavoidance at high intensities. It thus seems that the Bolwig organs and class IV dendritic arborization neurons operate in different light intensity regimes: Bolwig organs are tuned to low light, whereas class IV dendritic arborization neurons, required in low light, are the primary sensors at high intensities.

Careful examination of *ppk-GAL4* revealed additional expression in four mouth hook neurons, but not in the central nervous system (Supplementary Fig. 23a–d). Laser ablation of these four neurons in the *GMR-Hid* background had no effect on light avoidance behaviour (Supplementary Fig. 23e). Therefore, the class IV dendritic arborization neurons in the body wall are the ones important for the light avoidance behaviour.

*Pickpocket*, a Degenerin/Epithelial sodium Channel (DEG/ENaC) family member specifically expressed in class IV dendritic arborization neurons<sup>24,38</sup> (Fig. 3c), has been implicated in locomotion control<sup>39–41</sup>.



**Figure 5 | Class IV dendritic arborization neurons are the extra-ocular photoreceptors that contribute to light avoidance.** **a**, **b**, Examples of larvae with either class IV dendritic arborization neurons ablated (**a**) or both Bolwig organs and class IV dendritic arborization neurons ablated (**b**) that failed to respond to white light (0.57 mW mm<sup>-2</sup>) applied from 0 to 5 s (dotted circle). Arrow indicates locomotion direction. Scale bar, 1 mm (**a**, **b**), shown at -2 s. **c–g**, Percentage of animals avoiding white light of different intensities (in mW mm<sup>-2</sup>: **c**, 0.088; **d**, 0.24; **e**, 0.57; **f**, 1.0; **g**, 1.67). Wild-type larvae, Bolwig-organ-ablated larvae (*GMR-Hid*), larvae with class IV dendritic arborization neurons ablated (*ppk-GAL4*; *UAS-Hid,rpr*) and larvae with both ablated (*UAS-Hid,rpr*; *GMR-Hid*; *ppk-GAL4*) were examined. **h**, Percentage of Bolwig-organ-ablated animals avoiding 0.25 mW mm<sup>-2</sup> 525 nm green light when class IV dendritic arborization neurons express ChR2 with or without dietary retinal. **i**, Percentage of animals avoiding white light at 1 mW mm<sup>-2</sup>. For **c–i**, controls are black bars. Twenty to forty animals were tested for each condition; \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001; two-tailed Fisher exact test. ChR2, channelrhodopsin-2; *rpr*, *reaper*; NS, not significant.

However, nose-touch experiments<sup>42</sup> revealed that larvae lacking class IV dendritic arborization neurons responded normally to gentle touch by retracting or turning away their heads (Supplementary Fig. 24). Moreover, direct recording of class IV dendritic arborization neurons in *ppk* mutant larvae revealed no defect in light response (Supplementary Fig. 15b). These results demonstrate that reduced light avoidance in class IV dendritic-arborization-ablated larvae is not due to non-specific effects.

To probe sufficiency, we expressed channelrhodopsin-2 (ChR2), a retinal-dependent cation channel gated by light from ultraviolet to green<sup>43–45</sup>, specifically in class IV dendritic arborization neurons. ChR2 conferred green light sensitivity to dendritic arborization neurons from larvae fed with retinal (Supplementary Fig. 25), as well as robust avoidance of green light of retinal-fed larvae without Bolwig organs (Fig. 5h). Thus, activation of class IV dendritic arborization neurons is sufficient to induce avoidance.

With or without Bolwig organs, *TrpA1* mutant larvae showed deficient avoidance of 1 mW mm<sup>-2</sup> white light (Fig. 5i). Moreover, reducing *TrpA1* expression in class IV dendritic arborization neurons by RNAi was sufficient to abolish the light avoidance behaviour in animals without Bolwig organs (Supplementary Fig. 26). Together, our physiological and behavioural studies indicate that a light transduction pathway involving *TrpA1* and *Gr28b* in class IV dendritic arborization neurons is necessary for light avoidance.

## Discussion

Extra-ocular photoreceptors, previously found in reptiles, birds, amphibians and fish, provide a good measure of ambient light luminance and serve mainly non-image-forming functions such as phototaxis, circadian photo-entrainment, pupal reflex, shadow reaction and magnetic orientation<sup>1–7</sup>. Usually, these extra-ocular photoreceptors have much lower light sensitivity and slower kinetics than ocular photoreceptors<sup>3</sup>.

*Drosophila* larvae have primitive eye structures, the Bolwig organs, which control avoidance of dim light<sup>12</sup>. Here we report that the class IV dendritic arborization neurons, previously implicated in mechanosensory response and motion control<sup>38–41,46</sup>, are surprisingly also photoreceptors. Our behavioural analysis suggests that Bolwig organs and class IV dendritic arborization neurons have different regimes of light sensing in acute photoavoidance. Bolwig organs, packed with photopigments<sup>47</sup>, are preferentially required for avoidance of low light. Class IV dendritic arborization neurons, which also contribute to low light avoidance, are the primary sensors at sunlight-level intensities.

This organization ensures that larvae can detect the full range of ambient light intensities, from dim to strong.

Class IV dendritic arborization neurons have the intrinsic ability to sense light, even after isolation in culture, and their dendrites are capable of sensing light (Fig. 3 and Supplementary Fig. 10). Importantly, the dendrites of class IV dendritic arborization neurons have complete and non-redundant coverage of the body wall (Fig. 3c), allowing animals to perceive illumination of any body part, and initiate an appropriate behavioural response. Larvae spend much of the time with their heads digging into food, making their Bolwig organs on the head less likely to be exposed to light. Thus, the ability to sense light with sensory neurons tiling the body wall is critical for detection of exposure.

Class IV dendritic arborization neurons use a novel light transduction pathway. Like in *C. elegans*, a putative chemosensory G-protein-coupled receptor, *Gr28b*, is involved for phototransduction in class IV dendritic arborization neurons (Fig. 4b–e). *TrpA1* also is essential (Fig. 4f, g and Supplementary Fig. 18). *Drosophila* larval class IV dendritic arborization neurons may function as nociceptors<sup>46,48,49</sup>. They are required for thermal and mechanical nociception, and activation of class IV dendritic arborization neurons is sufficient to induce a behaviour pattern similar to nocifension<sup>46,48,49</sup>. Given that class IV dendritic arborization neurons are required for larvae to avoid harmful light stimuli, these neurons seem to be poised to alert the animal to a variety of adversities.

Our study has uncovered unexpected light-sensing machinery, which could be critical for foraging larvae to avoid harmful sunlight, desiccation and predation. By providing precedence for photoreceptors strategically placed away from the eyes, our finding of an array of class IV dendritic arborization neurons with elaborate dendrites tiling the entire body wall, and acting as light-sensing antennae, raises the question of whether other animals with eyes might also possess extra-ocular photoreceptors for more thorough light detection and behavioural response.

## METHODS SUMMARY

**Light avoidance assay.** Light avoidance was scored if the third instar larva reversed in direction or turned its head completely away from the 1.7-mm light spot on its head during the 5-s illumination. Two-tailed Fisher exact test (20–40 larvae per condition), \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001.

**Electrophysiology.** Action potentials were monitored via extracellular recordings from a third instar larval fillet with muscles removed, using an Axon 700B amplifier and pCLAMP 10 software.

**GCaMP3 imaging.** Third instar larval fillets were imaged on a Zeiss LS510 META confocal microscope with an Olympus  $\times 40/0.8$  NA water immersion objective, with a 488-nm laser. GCaMP3 cDNA is available from AddGene.

**Cell culture.** Embryos homozygous for Canton S (Cs); *UAS-GCaMP3*; *ppk-GAL4*, *UAS-RFP* (for class IV dendritic arborization neuron culture) or Cs; *UAS-GCaMP3*; *19-12-GAL4*, *UAS-RFP* (for class III dendritic arborization neuron culture) were used for culture<sup>22,23</sup>.

**MARCM analysis.** We recorded from class IV dendritic arborization neuron clones marked with GFP for lacking *TrpA1* (ref. 36).

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 8 March; accepted 13 October 2010.

Published online 10 November 2010.

1. Fu, Y., Liao, H. W., Do, M. T. & Yau, K. W. Non-image-forming ocular photoreception in vertebrates. *Curr. Opin. Neurobiol.* **15**, 415–422 (2005).
2. Yau, K. W. & Hardie, R. C. Phototransduction motifs and variations. *Cell* **139**, 246–264 (2009).
3. Steven, D. M. The dermal light sense. *Biol. Rev. Camb. Philos. Soc.* **38**, 204–240 (1963).
4. Millott, N. The dermal light sense. *Symp. Zool. Soc. Lond.* **23**, 1–36 (1968).
5. Yoshida, M. Extraocular photoreception. In *Handbook of Sensory Physiology* Vol. 7/6A, 581–640 (Springer, 1979).
6. Halford, S. et al. VA opsin-based photoreceptors in the hypothalamus of birds. *Curr. Biol.* **19**, 1396–1402 (2009).
7. Phillips, J. B., Deutschlander, M. E., Freake, M. J. & Borland, S. C. The role of extraocular photoreceptors in newt magnetic compass orientation: parallels between light-dependent magnetoreception and polarized light detection in vertebrates. *J. Exp. Biol.* **204**, 2543–2552 (2001).
8. Ward, A., Liu, J., Feng, Z. & Xu, X. Z. Light-sensitive neurons and channels mediate phototaxis in *C. elegans*. *Nature Neurosci.* **11**, 916–922 (2008).
9. Edwards, S. L. et al. A novel molecular solution for ultraviolet light detection in *Caenorhabditis elegans*. *PLoS Biol.* **6**, e198 (2008).
10. Liu, J. et al. *C. elegans* phototransduction requires a G protein-dependent cGMP pathway and a taste receptor homolog. *Nature Neurosci.* **13**, 715–722 (2010).
11. Sawin-McCormack, E. P., Sokolowski, M. B. & Campos, A. R. Characterization and genetic analysis of *Drosophila melanogaster* photobehavior during larval development. *J. Neurogenet.* **10**, 119–135 (1995).
12. Mazzoni, E. O., Desplan, C. & Blau, J. Circadian pacemaker neurons transmit and modulate visual information to control a rapid behavioral response. *Neuron* **45**, 293–300 (2005).
13. Sprecher, S. G. & Desplan, C. Switch of rhodopsin expression in terminally differentiated *Drosophila* sensory neurons. *Nature* **454**, 533–537 (2008).
14. Willson, R. C., Gulkis, S., Janssen, M., Hudson, H. S. & Chapman, G. A. Observations of solar irradiance variability. *Science* **211**, 700–702 (1981).
15. Grether, M. E., Abrams, J. M., Agapite, J., White, K. & Steller, H. The head involution defective gene of *Drosophila melanogaster* functions in programmed cell death. *Genes Dev.* **9**, 1694–1708 (1995).
16. Hay, B. A., Maile, R. & Rubin, G. M. P element insertion-dependent gene activation in the *Drosophila* eye. *Proc. Natl Acad. Sci. USA* **94**, 5195–5200 (1997).
17. Nakai, J., Ohkura, M. & Imoto, K. A high signal-to-noise  $\text{Ca}^{2+}$  probe composed of a single green fluorescent protein. *Nature Biotechnol.* **19**, 137–141 (2001).
18. Wang, J. W., Wong, A. M., Flores, J., Vossahl, L. B. & Axel, R. Two-photon calcium imaging reveals an odor-evoked map of activity in the fly brain. *Cell* **112**, 271–282 (2003).
19. Tian, L. et al. Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nature Methods* **6**, 875–881 (2009).
20. Ueda, A. & Wu, C. F. Effects of hyperkinetic, a  $\beta$  subunit of *Shaker* voltage-dependent  $\text{K}^{+}$  channels, on the oxidation state of presynaptic nerve terminals. *J. Neurogenet.* **22**, 103–115 (2008).
21. Auld, V. J., Fetter, R. D., Broadie, K. & Goodman, C. S. Gliotactin, a novel transmembrane protein on peripheral glia, is required to form the blood-nerve barrier in *Drosophila*. *Cell* **81**, 757–767 (1995).
22. Saito, M. & Wu, C. F. Expression of ion channels and mutational effects in giant *Drosophila* neurons differentiated from cell division-arrested embryonic neuroblasts. *J. Neurosci.* **11**, 2135–2150 (1991).
23. Bai, J., Sepp, K. J. & Perrimon, N. Culture of *Drosophila* primary cells dissociated from gastrula embryos and their use in RNAi screening. *Nature Protocols* **4**, 1502–1512 (2009).
24. Grueber, W. B. et al. Projections of *Drosophila* multidendritic neurons in the central nervous system: links with peripheral dendrite morphology. *Development* **134**, 55–64 (2007).
25. Grueber, W. B., Jan, L. Y. & Jan, Y. N. Tiling of the *Drosophila* epidermis by multidendritic sensory neurons. *Development* **129**, 2867–2878 (2002).
26. Grueber, W. B., Ye, B., Moore, A. W., Jan, L. Y. & Jan, Y. N. Dendrites of distinct classes of *Drosophila* sensory neurons show different capacities for homotypic repulsion. *Curr. Biol.* **13**, 618–626 (2003).
27. O'Tousa, J. E. et al. The *Drosophila ninaE* gene encodes an opsin. *Cell* **40**, 839–850 (1985).
28. Dolezelova, E., Dolezel, D. & Hall, J. C. Rhythm defects caused by newly engineered null mutations in *Drosophila's cryptochrome* gene. *Genetics* **177**, 329–345 (2007).
29. Montell, C. Visual transduction in *Drosophila*. *Annu. Rev. Cell Dev. Biol.* **15**, 231–268 (1999).
30. Thorne, N. & Amrein, H. Atypical expression of *Drosophila* gustatory receptor genes in sensory and central neurons. *J. Comp. Neurol.* **506**, 548–568 (2008).
31. Montell, C., Jones, K., Hafen, E. & Rubin, G. Rescue of the *Drosophila* phototransduction mutation *trp* by germline transformation. *Science* **230**, 1040–1043 (1985).
32. Rosenzweig, M. et al. The *Drosophila* ortholog of vertebrate TRPA1 regulates thermotaxis. *Genes Dev.* **19**, 419–424 (2005).
33. Hamada, F. N. et al. An internal thermal sensor controlling temperature preference in *Drosophila*. *Nature* **454**, 217–220 (2008).
34. Kwon, Y., Shim, H. S., Wang, X. & Montell, C. Control of thermotactic behavior via coupling of a TRP channel to a phospholipase C signaling cascade. *Nature Neurosci.* **11**, 871–873 (2008).
35. Kang, K. et al. Analysis of *Drosophila* TRPA1 reveals an ancient origin for human chemical nociception. *Nature* **464**, 597–600 (2010).
36. Lee, T. & Luo, L. Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis. *Neuron* **22**, 451–461 (1999).
37. White, K. et al. Genetic control of programmed cell death in *Drosophila*. *Science* **264**, 677–683 (1994).
38. Adams, C. M. et al. Ripped pocket and pickpocket, novel *Drosophila* DEG/ENAC subunits expressed in early development and in mechanosensory neurons. *J. Cell Biol.* **140**, 143–152 (1998).
39. Ainsley, J. A., Kim, M. J., Wegman, L. J., Pettus, J. M. & Johnson, W. A. Sensory mechanisms controlling the timing of larval developmental and behavioral transitions require the *Drosophila* DEG/ENAC subunit, Pickpocket1. *Dev. Biol.* **322**, 46–55 (2008).
40. Ainsley, J. A. et al. Enhanced locomotion caused by loss of the *Drosophila* DEG/ENAC protein Pickpocket1. *Curr. Biol.* **13**, 1557–1563 (2003).
41. Xu, K. et al. The *fragile X-related* gene affects the crawling behavior of *Drosophila* larvae by regulating the mRNA level of the DEG/ENAC protein Pickpocket1. *Curr. Biol.* **14**, 1025–1034 (2004).
42. Kernan, M., Cowan, D. & Zuker, C. Genetic dissection of mechanosensory transduction: mechanoreception-defective mutations of *Drosophila*. *Neuron* **12**, 1195–1206 (1994).
43. Nagel, G. et al. Channelrhodopsin-2, a directly light-gated cation-selective membrane channel. *Proc. Natl Acad. Sci. USA* **100**, 13940–13945 (2003).
44. Schroll, C. et al. Light-induced activation of distinct modulatory neurons triggers appetitive or aversive learning in *Drosophila* larvae. *Curr. Biol.* **16**, 1741–1747 (2006).
45. Zhang, F., Aravanis, A. M., Adamantidis, A., de Lecea, L. & Deisseroth, K. Circuit-breakers: optical technologies for probing neural signals and systems. *Nature Rev. Neurosci.* **8**, 577–581 (2007).
46. Zhong, L., Hwang, R. Y. & Tracey, W. D. Pickpocket is a DEG/ENAC protein required for mechanical nociception in *Drosophila* larvae. *Curr. Biol.* **20**, 429–434 (2010).
47. Pollock, J. A. & Benzer, S. Transcript localization of four opsin genes in the three visual organs of *Drosophila*; RH2 is ocellus specific. *Nature* **333**, 779–782 (1988).
48. Tracey, W. D., Jr, Wilson, R. I., Laurent, G. & Benzer, S. *Painless*, a *Drosophila* gene essential for nociception. *Cell* **113**, 261–273 (2003).
49. Hwang, R. Y. et al. Nociceptive neurons protect *Drosophila* larvae from parasitoid wasps. *Curr. Biol.* **17**, 2105–2116 (2007).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank P. Garrity, C. Desplan, J. Blau, C. Montell, J. Hall, H. Amrein, L. Tian, S. Younger and S. Zhu for fly stocks and reagents; T. Jin for technical support; C. Han for generating whole larval images; H. H. Lee for collaboration to identify the 19-12-GAL4 and 21-7-GAL4 lines; R. Yang, H. H. Lee, B. Ye, J. Parrish, P. Soba, B. Schroeder and J. Bagley for discussions and advice; B. Ye, R. Yang, W. Ge and J. Berg for critical reading of the manuscript; and Jan laboratory members for discussions. Y.X. was a recipient of a Long-term Fellowship from the Human Frontier Science Program. N.V. is supported by Deutsche Forschungsgemeinschaft. This work is supported by a NIH grant (2R37NS040929) to Y.N.J. Y.X. and Q.Y. are associates, L.Y.J. and Y.N.J. are investigators of the Howard Hughes Medical Institute. L.L.L. is supported by the Howard Hughes Medical Institute, Janelia Farm Campus.

**Author Contributions** Y.X. designed and carried out the experiments and analysed the data; Q.Y. characterized molecular information of Gr28b, rhodopsin and cryptochrome. L.L.L. created GCaMP3 and did the bioinformatic analyses of Gr28b; N.V. cleaned up the *Rh3<sup>1</sup>* and *Rh4<sup>1</sup>* mutants; Y.N.J. helped to design the experiments and supervised the work; Y.X., L.L.L., L.Y.J. and Y.N.J. wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to Y.N.J. ([yuhnung.jan@ucsf.edu](mailto:yuhnung.jan@ucsf.edu)).



## METHODS

**Fly stocks.** The following fly strains were used: (1) *Cs*; (2) *Cs*; *GMR-Hid*; (3) *Cs*; *ppk-GAL4*, *UAS-Tomato*; (4) *Cs*; *ppk-GAL4*, *UAS-mCD8::GFP*; (5) *Cs*; *ppk-GAL4*, *UAS-mCD8::RFP*; (6) *Cs*; *TrpA1*; (7) *w*; *UAS-TrpA1 RNAi*; (8) *Cs*; *UAS-Dicer*; *ppk-GAL4*, *UAS-mCD8::GFP*; (9) *Cs*; *21-7-GAL4*, *UAS-GCaMP3*; (10) *Cs*; *UAS-GCaMP3*; *ppk-GAL4*, *UAS-mCD8::RFP*; (11) *Cs*; *UAS-GCaMP3*; *19-12-GAL4*, *UAS-mCD8::RFP*; (12) *elav-GAL4*, *hsFLP*, *UAS-mCD8::GFP*; *tub-GAL80*, *FRT<sup>2A</sup>*; (13) *w*; *TrpA1*, *FRT<sup>2A</sup>/TM6B*, *Tb*; (14) *Cs*; *MiET103888* (Bloomington stock centre no. 24190); (15) *Cs*; *PBac01884* (Bloomington stock centre no. 10743); (16) *w*; *UAS-Gr28b RNAi* (VDRC stock centre no. kk101727); (17) *Cs*; *cry<sup>01</sup>* (ref. 28); (18) *Cs*; *Rh3<sup>1</sup>*; (19) *Cs*; *Rh4<sup>1</sup>*; (20) *Cs*; *Rh5<sup>2</sup>*; (21) *Cs*; *Rh6<sup>1</sup>*; (22) *ninaE<sup>17</sup>*; (23) *norpA<sup>P24</sup>*; (24) *Df(2L)Exel7031*; (25) *yw*, *UAS-Hid*, *rpr*; (26) *Cs*; *Channelrhodopsin-2*; (27) *UAS-TrpA1*; (28) *Cs*; *trpl<sup>302</sup>*; (29) *Cs*; *painless<sup>1</sup>*.

**Light avoidance assay.** Animals were raised at 25 °C in an incubator with 12 h light/dark cycles and humidity control (Darwin Chamber Company). Ninety-six hours after egg laying (AEL), third instar larvae were gently picked up from the vial, washed twice with PBS and transferred to a 100-mm Petri dish with fresh 2% agarose. Excessive water was removed from the animals. Animals were allowed to rest on the plate for at least 3 min before testing. Only animals making straight forward movement were selected for the assay. Each animal was tested once. The assay was carried out with a Stereo Microscope system (Leica M205FA). Unless otherwise specified, light was delivered from a 300 W xenon lamp (Sutter LB-LS/30) through a PLANAPO  $\times 1$  objective (Leica) at  $\times 160$  magnification, yielding a light spot of 1.7 mm in diameter. To direct the light to the animal's head, the plate was manually moved so that only the head appeared in the field of view. An avoidance response was scored when animals stopped forward movement during the 5-s light illumination by either initiating backward movement or turning their heads completely away from the light spot. The 5-s light illumination was controlled by a shutter (Sutter Instruments) in the xenon lamp house triggered by an external stimulator (Grass s88). The xenon lamp has a light intensity spectrum similar to sunlight. The background light for visualizing animals was filtered to red, to which they are insensitive (Lee filter no. 027, medium red), and the entire event was recorded through a lens (Fujinon, 25 mm 1:1.4) to a CCD camera (Qimaging Rolera XR) at 6 frames per second. The camera was mounted on a tripod and placed on the side of the Petri dish, with the front of the lens covered with red filters (Lee filter no. 027, medium red) to avoid overexposure. For violet, blue, green or red light, the band-pass excitation filter (in nm:  $402 \pm 7.5$ ,  $470 \pm 20$ ,  $525 \pm 25$ ,  $620 \pm 30$ ) was placed in the xenon lamp house, and a filter set with empty excitation filter was placed into the Leica scope. White light illumination was achieved the same way except that no excitation filter was placed in the xenon lamp house. For 360 nm illumination, a HXP-120 light source (VisiTron Systems) was used, and a  $360 \pm 20$  nm excitation filter set was placed into the Leica scope. The light intensity at  $\times 160$  magnification was measured by a radiometric sensor head (Newport 818P-001-12) coupled with a power meter (Newport 1918-C). Because liquid light guides connecting the light source and the microscope ensured the uniformity of light, the light intensity was calculated by dividing the measured light intensity over area ( $2.27 \text{ mm}^2$ ). Temperature changes associated with light illumination were measured with an IT-24P thermocouple probe coupled with a BAT-10 thermometer (Physitemp). A red filter (Lee filter no. 027, medium red) was used to cover the eyepieces of the microscope to protect experimenters' eyes from strong light. Light avoidance of animals expressing channelrhodopsin-2 (ChR2) to  $0.25 \text{ mW mm}^{-2}$  green light was done the same way with the exception that eggs were laid and allowed to develop to third instar in food medium supplemented with  $0.2 \text{ mM}$  retinal. Twenty to forty animals were tested in each condition and the percentage of positive responses was calculated. A two-tailed Fisher exact test was performed and statistical significance was assigned,  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ .

**Cell culture.** Dissociated cell cultures were prepared from early gastrulas of *Drosophila melanogaster*, as described previously<sup>22,23</sup>. Briefly, embryos were collected on grape agar plates with yeast paste and incubated for another 3.5 h at 25 °C. After removing yeast paste carefully, embryos were washed extensively with 500 ml sterile  $\text{H}_2\text{O}$ . To remove the chorionic membrane and sterilize the embryos, embryos were treated with bleach/90% EtOH (1:1 by volume, final concentration of sodium hypochlorite is 3%) solution for 1 min and then washed with 500 ml sterile  $\text{H}_2\text{O}$  to remove residual bleach and EtOH. After wash, embryos were homogenized in Schneider's medium supplemented with 5% FBS,  $0.2 \mu\text{g ml}^{-1}$  insulin, penicillin ( $50 \text{ units ml}^{-1}$ ) and streptomycin ( $50 \mu\text{g ml}^{-1}$ ), with a 15-ml Dounce homogenizer containing 6 ml medium. Three to five rotary up and down strokes were used to dissociate the cells. To remove undissociated clumps and large debris, dissociated cells were filtered through a  $40\text{-}\mu\text{m}$  cell strainer. Filtrate was collected in a 15-ml centrifuge tube, and cells were centrifuged at 2,000 r.p.m. for 5 min. Cells were washed in the medium and centrifuged again as described, followed by a final suspension in 10 ml medium.

Cells were plated on poly-L-lysine-treated coverglass, and allowed to develop at room temperature for 2–4 days before GCaMP3 imaging was performed. Embryos carrying *Cs*; *UAS-GCaMP3*; *ppk-GAL4*, *UAS-mCD8::RFP* were used for culturing class IV dendritic arborization neurons, and embryos carrying *Cs*; *UAS-GCaMP3*; *19-12-GAL4*, *UAS-mCD8::RFP* were used for culturing class III dendritic arborization neurons.

**Electrophysiology.** Fillets were made from 96-h-AEL third instar larvae with cuticle facing down in the external saline solution composed of (in mM): NaCl 120, KCl 3,  $\text{MgCl}_2$  4,  $\text{CaCl}_2$  1.5,  $\text{NaHCO}_3$  10, trehalose 10, glucose 10, TES 5, sucrose 10, HEPES 10. Osmolality was  $305 \text{ mOsm kg}^{-1}$  and final pH = 7.25. Muscles covering the neurons of interest were gently digested by proteinase (Sigma). During the enzymatic treatment, a laminar flow of external saline solution was turned on to remove excess enzyme. No detectable difference in dendrite morphology of class IV dendritic arborization neurons was observed before and after muscle digestion, indicating that neurons were intact. An Olympus BX51WI microscope with a  $\times 40/0.8$  NA water immersion objective was used to obtain recordings with the help of IR-DIC optics and a CoolSNAP CCD (Photometrics). Recording pipettes were pulled with P-97 puller (Sutter instruments) from thin wall borosilicate glass (World Precision Instruments), filled with external saline solution, with a tip opening of  $5 \mu\text{m}$ . Gentle negative pressure was delivered to suck the soma to get good signal-to-noise ratio of recording traces. Recordings were performed with a Multiclamp 700B amplifier (Molecular Devices), and data were acquired with Digidata 1440A (Molecular Devices) and Clampex 10.0 software (Molecular Devices). Extracellular recordings of action potentials were obtained in voltage clamp mode with a holding potential of 0 mV, with a 2 kHz low-pass filter and sampled at 20 kHz. During recording, no background light illumination was applied. A 300-W xenon light source was connected to the microscope with a liquid light guide to provide light stimulation through a  $\times 40/0.8$  NA water-immersion lens, yielding an evenly illuminated light spot with  $600 \mu\text{m}$  diameter, which covered the entire class IV dendritic arborization neuron. Dendritic illumination was achieved by decreasing the field diaphragm to cover about 50% of dendrites (Fig. 3d), without illuminating the soma. Light intensity was measured the same way as in the behaviour assay and intensity density was calculated. Neutral density filters (Chroma) were used to reduce light intensity to generate dose-response curves. The duration (5 s) of light illumination was controlled by a shutter in the xenon lamp house triggered by Digidata 1440A (Molecular Devices). Band-pass excitation filters (Chroma) were used to select light wavelength: they were (in nm)  $340 \pm 10$ ,  $380 \pm 10$ ,  $402 \pm 7.5$ ,  $470 \pm 20$ ,  $525 \pm 25$ ,  $620 \pm 30$ . White light illumination was achieved the same way with the exception that no excitation filter was placed in the xenon lamp house. For each recording trace, average frequency during the 5 s immediately before light exposure was used as control. Five-second light stimulation was controlled by a TTL-triggered shutter (Sutter Instruments) in the xenon lamp house. For latency analysis, only neurons with low spontaneous firing were included for recording, and latency was defined as the time between onset of light and onset of burst firing. To record temperature-induced firing change, pre-heated solution was perfused into the recording chamber. Temperature was monitored by the thermal probe connected with the thermometer (Warner TC-324B). mSIRK (EMD bioscience) or *L-cis*-diltiazem (Sigma) was incubated in the recording chamber for 30 min before recording. A two-tailed paired or unpaired *t*-test, or one-way ANOVA, followed by the Bonferroni multiple comparison test, was performed and statistical significance was assigned,  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ .

**GCaMP3 imaging experiments.** Homozygous *Cs*; *21-7-GAL4*, *UAS-GCaMP3* animals were used for imaging. The *21-7* promoter drives Gal4 expression in all peripheral nervous system sensory neurons in the larval body wall except the chordotonal organ (H. H. Lee, Y.X., L.Y.J. and Y.N.J., unpublished data). Fillet preparation was the same as the one used in recording with the exception that the cuticle was facing up to mimic the orientation of larvae in receiving natural light. No enzymatic digestion was performed. Data were collected on a Zeiss LS510 META confocal microscope with an Olympus  $\times 40/0.8$  NA water immersion objective. GCaMP3 fluorescence was excited with a 488-nm laser<sup>18,19</sup>. Laser scanning by itself didn't activate class IV dendritic arborization neurons, as evidenced by the flat baseline of GCaMP3 signals during the 0–60 s control period (Fig. 2e). The images were acquired at  $512 \times 512$  pixels at 12-bit dynamic range. The duration of 5 s of light stimulation ( $91$ ,  $68$  and  $96 \text{ mW mm}^{-2}$  for 365 nm ultraviolet, 470 nm blue and 546 nm green light, respectively) was controlled by manually switching the filter cube from laser-scanning position to epifluorescence position. Two seconds of light stimulation elicited similar results as five seconds (data not shown). Average GCaMP3 signals from 60 s before light stimulation was taken as  $F_0$ , and  $\Delta F/F_0$  was calculated for each data point. GCaMP3 signals from the soma were analysed, although axons and dendrites also showed responses. As a control, *Cs*; *UAS-GCaMP3*; *ppk-GAL4*, *UAS-mCD8::RFP* animals

were raised to look for nonspecific effects of excitation. In these animals, we did not detect any change in RFP signals in response to light (data not shown).

For GCaMP3 imaging of cultured neurons, class IV or class III dendritic arborization neurons were identified by the co-expressed RFP signals. Light stimulation was carried out as described above. For AITC stimulation, an equal volume of 200  $\mu$ M AITC was manually applied to the chamber, making the final concentration 100  $\mu$ M.

**MARCM analysis.** MARCM analysis was carried out as described previously<sup>36</sup>. *TrpA1* mutant class IV dendritic arborization neurons with GFP signals were

selected for recording. For *TrpA1*<sup>+/-</sup> heterozygous neurons without GFP expression, class IV dendritic arborization neurons were identified by location.

**Nose-touch assay.** Assay was performed as described previously<sup>42</sup>. Briefly, the larvae were touched with an eyebrow hair affixed to the tip of a dissecting needle. The scoring system is as follows: 0 = no response to touch; 1 = a response of pausing mouth hook movement; 2 = responding by withdrawing the anterior or turning away from the touch; 3 = a single reverse peristaltic wave away from the touch; and 4 = multiple peristaltic waves away from the touch. A two-tailed Fisher exact test was performed.

# TRIM24 links a non-canonical histone signature to breast cancer

Wen-Wei Tsai<sup>1\*</sup>, Zhanxin Wang<sup>3\*</sup>, Teresa T. Yiu<sup>1,2</sup>, Kadir C. Akdemir<sup>2,5</sup>, Weiya Xia<sup>4</sup>, Stefan Winter<sup>6</sup>, Cheng-Yu Tsai<sup>7</sup>, Xiaobing Shi<sup>1,2</sup>, Dirk Schwarzer<sup>8</sup>, William Plunkett<sup>7</sup>, Bruce Aronow<sup>9</sup>, Or Gozani<sup>10</sup>, Wolfgang Fischle<sup>6</sup>, Mien-Chie Hung<sup>4,11</sup>, Dinshaw J. Patel<sup>3</sup> & Michelle Craig Barton<sup>1,2</sup>

**Recognition of modified histone species by distinct structural domains within ‘reader’ proteins plays a critical role in the regulation of gene expression. Readers that simultaneously recognize histones with multiple marks allow transduction of complex chromatin modification patterns into specific biological outcomes. Here we report that chromatin regulator tripartite motif-containing 24 (TRIM24) functions in humans as a reader of dual histone marks by means of tandem plant homeodomain (PHD) and bromodomain (Bromo) regions. The three-dimensional structure of the PHD–Bromo region of TRIM24 revealed a single functional unit for combinatorial recognition of unmodified H3K4 (that is, histone H3 unmodified at lysine 4, H3K4me0) and acetylated H3K23 (histone H3 acetylated at lysine 23, H3K23ac) within the same histone tail. TRIM24 binds chromatin and oestrogen receptor to activate oestrogen-dependent genes associated with cellular proliferation and tumour development. Aberrant expression of TRIM24 negatively correlates with survival of breast cancer patients. The PHD–Bromo of TRIM24 provides a structural rationale for chromatin activation through a non-canonical histone signature, establishing a new route by which chromatin readers may influence cancer pathogenesis.**

Post-translational modifications of histones occur in combinations that must be faithfully translated by effector proteins, or histone readers<sup>1–4</sup>. The lexicon of histone modifications may be highly context-dependent, influenced by inductive signalling, cellular milieu and target gene status<sup>4</sup>. Misinterpretation or imbalance in this hierarchical arrangement has dire consequences for cellular homeostasis, leading to developmental problems, hereditary disease or tumour development<sup>5</sup>. Linked histone reader modules, such as tandem PHD finger and bromodomain, occur frequently in proteins that interact with histone, but little is known about their mechanisms of action. Combinatorial readout of histone post-translational modifications (PTMs) may enhance binding between spatially separated histone marks, or even create communication links between domains or members of the complex<sup>3</sup>. Individually, proteins with bromodomains—for example, TAF1 and BDF1—associate with acetylated lysines with broad specificity<sup>6,7</sup>, while PHD-containing proteins are less predictable in their interactions<sup>1–4</sup>. The PHD fingers of BHC80 and AIRE interact with unmethylated H3K4 (H3K4me0)<sup>8,9</sup>, while other previously reported PHD finger domains bind methylated proteins as modifiers of histones or as subunits of chromatin-remodelling, co-activator or co-repressor complexes<sup>1–4,10–17</sup>.

PHD-finger proteins and their dysregulation are linked to a broad spectrum of human diseases, underscoring an essential role in homeostasis<sup>5</sup>. Recently, aberrant localization of a JARID1A PHD finger-fusion protein was shown as directly causal in transformation and development of haematopoietic malignancy, which is a process requiring fusion protein recognition of H3K4me3 via the JARID1A PHD finger<sup>18</sup>. Here, we present evidence that a multi-functional

protein, TRIM24, which is an E3-ubiquitin ligase that targets p53 (ref. 19) and is broadly associated with chromatin silencing<sup>20</sup>, relies on PHD-Bromo to recognize specific, combinatorial histone modifications and activate oestrogen-dependent genes associated with cellular proliferation and tumour development. Genome-wide analysis of chromatin interactions shows oestrogen-dependent binding of TRIM24 and oestrogen receptor  $\alpha$  (ER $\alpha$ ) at sites that paradoxically exhibit oestrogen-activated loss of H3K4me2 and gain of histone acetylation. Importantly, aberrant overexpression of TRIM24 in breast cancer patients is frequent and directly correlated with poor survival.

## TRIM24 PHD–Bromo binds amino-terminal H3 tail

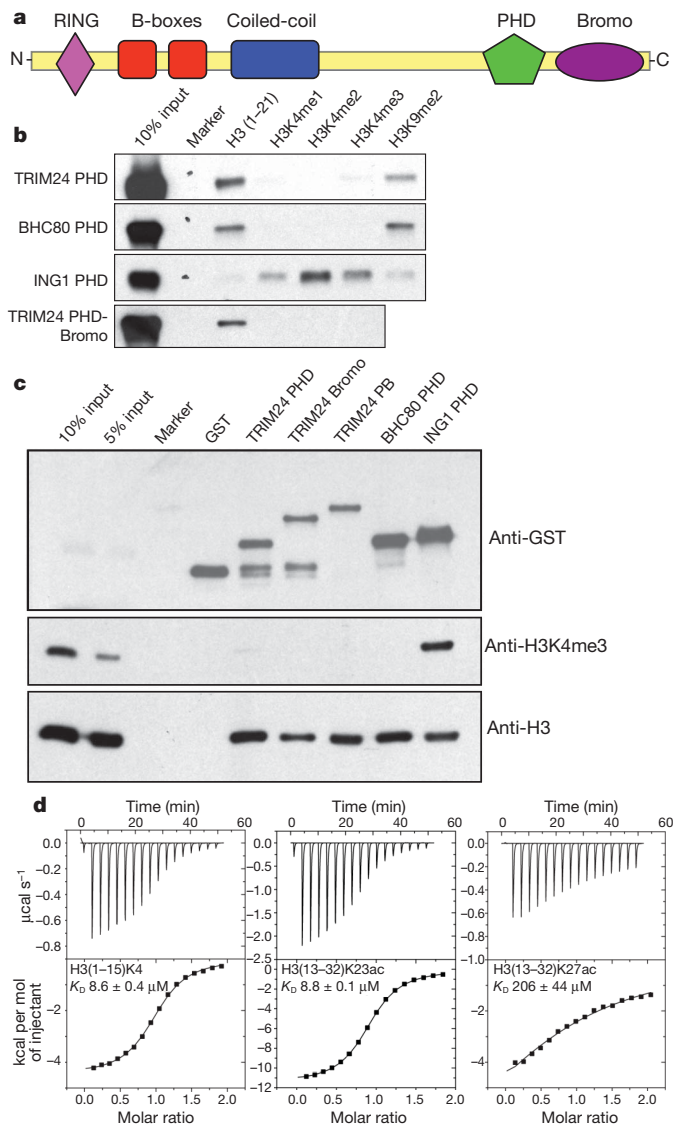
TRIM24 belongs to the TRIM/RBCC protein family, characterized by a conserved, N-terminal tripartite motif—namely, a RING domain, B-box zinc-fingers, and a coiled-coil region—as well as variable carboxy-terminal domains<sup>21,22</sup>. TRIM24 was originally identified as transcriptional intermediary factor (TIF) 1 $\alpha$ , a ligand-dependent, co-repressor of retinoic acid receptor that interacts with multiple nuclear receptors *in vitro* via an LXXLL motif<sup>23</sup>. In addition to its LXXLL motif and RING domain, TRIM24 has a C-terminal, PHD-Bromo (Fig. 1a), which probably recognizes histones or non-histone proteins with specific combinations of post-translational modifications.

Protein sequence alignment of the PHD fingers of TRIM24 and BHC80 with ING1, a PHD domain that recognizes H3K4me3 (refs 24,25), showed TRIM24 as highly similar to BHC80 with conservation of residues critical for BHC80–H3K4me0 interactions<sup>8</sup>

<sup>1</sup>Department of Biochemistry and Molecular Biology, Program in Genes and Development, Graduate School of Biomedical Sciences, University of Texas M.D. Anderson Cancer Center, Houston, Texas 77030, USA. <sup>2</sup>Centers for Cancer Epigenetics and Stem Cell and Developmental Biology, University of Texas M.D. Anderson Cancer Center, Houston, Texas 77030, USA. <sup>3</sup>Structural Biology Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. <sup>4</sup>Department of Molecular and Cellular Oncology, University of Texas M.D. Anderson Cancer Center, Houston, Texas 77030, USA. <sup>5</sup>Department of Biostatistics and Bioinformatics, University of Texas M.D. Anderson Cancer Center, Houston, Texas 77030, USA. <sup>6</sup>Laboratory of Chromatin Biochemistry, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Goettingen, Germany. <sup>7</sup>Department of Experimental Therapeutics, University of Texas M.D. Anderson Cancer Center, Houston, Texas 77030, USA. <sup>8</sup>Department of Chemical Biology/Protein Chemistry, Leibniz-Institut für Molekulare Pharmakologie (FMP), Robert-Rössle-Strasse 10, 13125 Berlin, Germany. <sup>9</sup>Computational Medicine Center, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio 45229, USA. <sup>10</sup>Department of Biological Sciences, Stanford University, Stanford, California 94305, USA. <sup>11</sup>Center for Molecular Medicine and Graduate Institute of Cancer Biology, China Medical University and Hospital, Taichung 404, Taiwan.

\*These authors contributed equally to this work.





**Figure 1 | TRIM24 PHD finger interacts with unmethylated H3K4.**

**a**, Diagram of TRIM24 protein domains. **b**, Biotinylated peptide pull-downs: recombinant PHD fingers and histone peptides. **c**, GST-pull-downs: recombinant proteins and native histone proteins, PHD-Bromo (PB). **d**, ITC titration: binding of TRIM24 PHD-Bromo with histone peptides.

(Supplementary Fig. 1a). Accordingly, we found that full-length TRIM24 interacts with histone proteins specifically through its PHD-Bromo (Supplementary Fig. 1b). Binding of the TRIM24 PHD-Bromo to histone peptide arrays occurs at unmodified H3 (residues 1–21), methylated H3K9 (H3K9me) and acetylated H3K9/K14 peptides, but not methylated H3K4 residues (Supplementary Fig. 1c). Similarly, TRIM24 PHD finger and PHD-Bromo bind unmodified histone H3 (residues 1–21) but not methylated H3K4, similar to BHC80 but unlike ING1, which preferentially binds to H3K4me peptides (Fig. 1b and Supplementary Fig. 1d). Glutathione S-transferase (GST)-pull-down assays with native histones confirmed that TRIM24 PHD finger, bromodomain, PHD-Bromo and the BHC80 PHD fail to bind to native histone H3 with K4 trimethylation (H3K4me3) but tolerate H3K9me2 modification (Fig. 1c and Supplementary Fig. 1e). Isothermal titration calorimetry (ITC)-based binding assays established that the PHD-Bromo binds unmodified H3(1–15)K4 with a dissociation constant,  $K_D$ , of 8.6  $\mu$ M, while methylation of H3K4 greatly decreases binding affinity of TRIM24 and H3 peptides (Fig. 1d and Supplementary Table 2). These results suggest that TRIM24 PHD-Bromo interacts with the N-terminal tail of histone

H3, but that specific PTMs, for example, methylation of H3K4, interfere with this interaction.

## Structural basis of H3 readout by TRIM24

We have determined the three-dimensional crystal structure of the PHD-linker-Bromo segment (residues 824–1006) of human TRIM24 in free and histone peptide bound states. The overall structure of TRIM24 PHD-Bromo in the free state demonstrates that PHD and bromodomain interact extensively and form an integrated structural unit (747  $\text{\AA}^2$  of contact surface), connected by a long linker and stabilized by a network of hydrogen bonding and hydrophobic interactions (Fig. 2a, Supplementary Fig. 2, Supplementary Table 1). The TRIM24 PHD finger residues 824–871 adopt the typical PHD finger ‘cross-braced’ topology stabilized by a pair of coordinated zinc ions, which together with residues 872–884 from the linker region form an extended TRIM24 PHD domain. The TRIM24 bromodomain adopts the typical left-handed four-helical bundle characteristic of other members of this family.

The 2.0  $\text{\AA}$  co-crystal structure of TRIM24 PHD-Bromo and unmodified H3(1–10)K4 peptide (Supplementary Table 1 and Supplementary Fig. 3a) showed that the first nine residues of bound H3 peptide are positioned within a surface groove of the PHD finger (Fig. 2b and Supplementary Fig. 3b). The R2 to Q5 segment of bound H3 peptide forms an anti-parallel  $\beta$ -sheet with the E837 to C840 segment of the PHD finger, while the T6 to K9 segment of bound H3 peptide contacts the N834 to G836 segment of the PHD finger. The side chain of R2 is hydrogen-bonded with the backbone carbonyl of C841. The side chain of C840 is positioned in-between the side chains of R2 and K4, with the C840W mutation losing its ability to bind unmodified H3K4 peptide ( $K_D > 400 \mu\text{M}$ , Supplementary Table 2 and Supplementary Fig. 4).

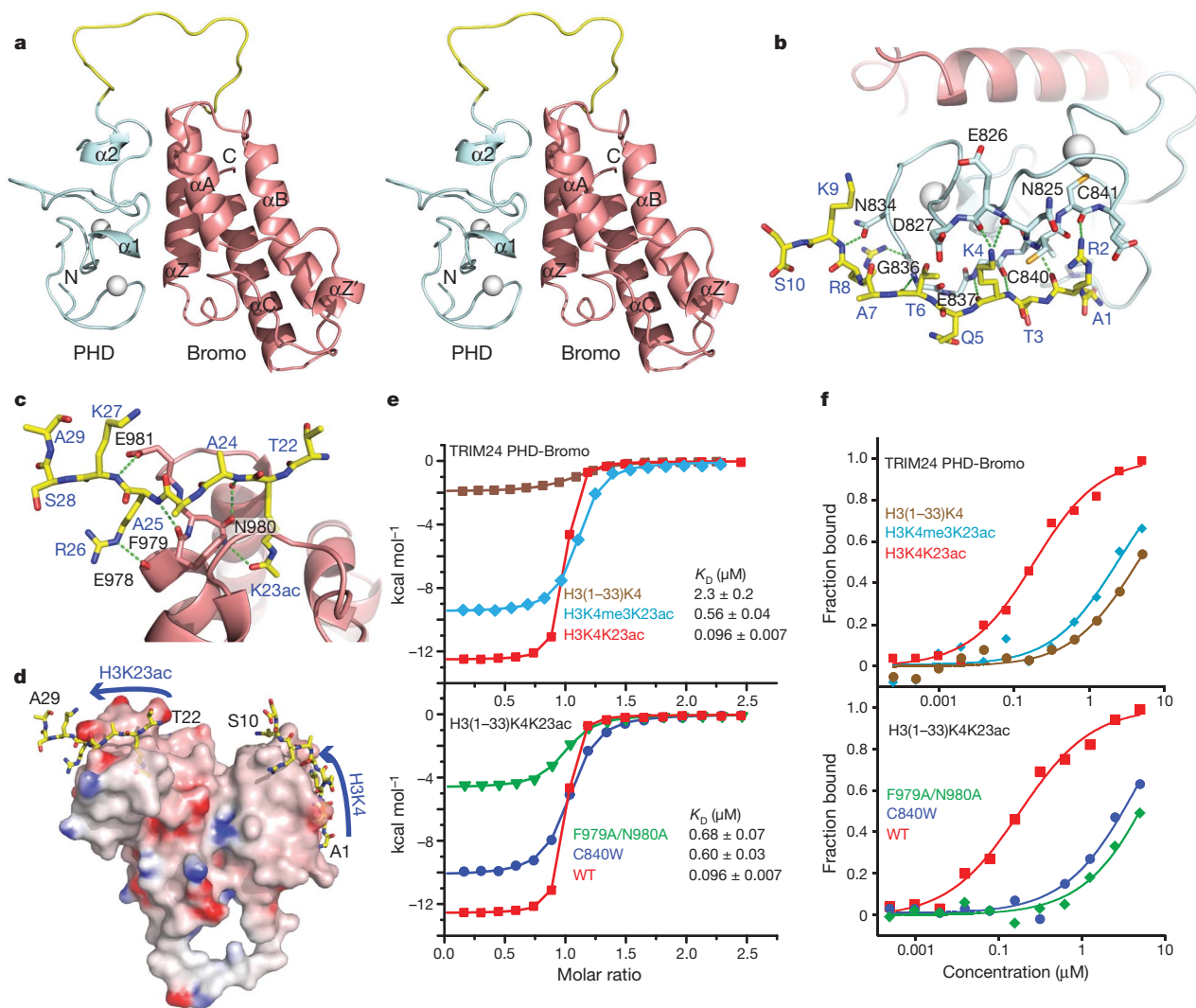
The unmodified lysine ammonium group of H3K4 forms two direct hydrogen bonds with backbone carbonyl oxygens of N825 and E826 (Fig. 2b). In addition, the proximally positioned D827 forms a stabilizing salt bridge with the unmodified lysine, consistent with the observation of impaired binding between D827A mutant and unmodified H3K4 peptide ( $K_D = 133 \mu\text{M}$ , Supplementary Table 2). Methylation of H3K4 would create steric clashes with residues lining the binding pocket, disrupt the salt bridge interaction with D827, and impair hydrogen bonding with N825 and E826, thereby providing a structural explanation for the unmodified H3K4 preference of TRIM24 PHD-Bromo.

## TRIM24 bromodomain is H3K23ac-specific

Both sequence and structure-based alignments indicate that TRIM24 bromodomain is an acetyllysine reader. Peptide pull-down assays and NMR titration measurements suggest that TRIM24 bromodomain interacts with H3 peptides with K23 or K27 acetylation and several acetylated H4 peptides (Supplementary Fig. 5a, b). ITC studies establish that TRIM24 PHD-Bromo specifically binds to the H3(13–32)K23ac peptide with a value of  $K_D$  (8.8  $\mu\text{M}$ ; Supplementary Table 2) comparable to tetra-acetylated H4 peptide and double bromodomain modules of TAF1 or BDF1.

We solved the 1.9  $\text{\AA}$  crystal structure of the complex of TRIM24 PHD-Bromo and H3(13–32)K23ac peptide (Supplementary Table 1 and Supplementary Fig. 6a). Residues 23–27 of the bound H3(13–32)K23ac peptide exhibit sequence-specific interactions with TRIM24 bromodomain (Fig. 2c and Supplementary Fig. 6b). The acetyllysine side chain forms a direct hydrogen bond with the side chain of conserved N980. Acetyllysine recognition constitutes the binding determinant, as double mutant F979A/N980A loses most of the binding affinity for the H3(13–32)K23ac peptide (Supplementary Table 2).

ITC studies establish that H3(1–20)K9ac, H3(1–19)K14ac and H3(13–32)K27ac bind non-specifically to the TRIM24 bromodomain ( $K_D \approx 200 \mu\text{M}$ ; Supplementary Table 2). The crystal structure of the complex of TRIM24 PHD-Bromo with H3(23–31)K27ac peptide (Supplementary Table 1) revealed a single intermolecular hydrogen



**Figure 2 | TRIM24 PHD-Bromo simultaneously binds H3K4me0 and acetylated histone lysines.** **a**, Stereo view of the crystal structure of TRIM24 PHD-Bromo in the free state. **b**, Detailed interactions between PHD of TRIM24 PHD-Bromo and H3(1-10)K4 peptide. **c**, Detailed interactions between bromodomain of TRIM24 PHD-Bromo and H3(22-29)K23ac peptide. **d**, Positioning of H3(1-10)K4 and H3(13-32)K23ac peptides on the surface of

TRIM24 PHD-Bromo based on structural information. **e**, **f**, ITC (**e**) or fluorescence polarization (FP) (**f**) based binding curves of wild-type (WT) or mutant forms of TRIM24 PHD-Bromo with H3(1-33) peptides bearing different combination of modifications. Insets,  $K_D$  values derived from ITC experiments.

bond between the side chains of K27ac and N980, while other histone residues did not show any direct intermolecular contacts with the bromodomain (Supplementary Fig. 7), consistent with weak binding affinity of H3(13-32)K27ac peptide. The structure of TRIM24 PHD-Bromo bound to H4(14-19)K16ac peptide containing the conserved interaction between K16ac and N980 side chains is shown in Supplementary Fig. 8.

The structures of TRIM24 PHD-Bromo complexes with acetyllysine-containing histone peptides show that acetyllysine invariably inserts into a pre-formed acetyllysine-binding pocket of the bromodomain. With the acetyllysine as an anchor, flanking residues determine sequence specificity of acetyllysine peptides for the TRIM24 bromodomain. The H3(13-32)K23ac peptide both fits better within the cleft between ZA and BC loops, and shows sequence-specific interactions with TRIM24 bromodomain spanning K23ac to K27, creating much higher affinity for the TRIM24 bromodomain than is shown by other acetyllysine-containing peptides.

### Combinatorial readout by TRIM24 PHD-Bromo

Superimposition of the above structures of complexes revealed that H3K4 and H3K23ac peptides are aligned in the same direction on the

surface of the TRIM24 PHD-Bromo (Fig. 2d). The distance between the C $\alpha$  of H3K9 and the C $\alpha$  of H3K23ac is 25.5 Å, which allows one H3 peptide containing both unmodified H3K4 and H3K23ac to simultaneously target the PHD and bromodomain binding sites on TRIM24 PHD-Bromo.

By contrast, H3K4 and H3K27ac (or H4K16ac) peptides are aligned in opposite directions on the surface of TRIM24 PHD-Bromo (Supplementary Fig. 9), which indicates that the TRIM24 PHD-Bromo requires two histone tails, either within a single nucleosome or from an adjacent pair of nucleosomes, to simultaneously bind H3K4 and H3K27ac (or H4K16ac).

To test the effect of combinatorial readout of TRIM24 PHD-Bromo on histone H3 bearing unmodified K4 and acetylated K23 dual marks, we synthesized longer H3(1-33) peptides bearing both unmodified K4 and acetylated K23 marks. For controls, we used H3(1-33)K4me3K23ac, as well as H3(1-33)K4 peptides that have only one effective histone mark for specific TRIM24 PHD-Bromo recognition. On the basis of ITC binding assays, TRIM24 PHD-Bromo showed an approximately 90-fold higher binding affinity for H3(1-33)K4K23ac peptide (Fig. 2e,  $K_D = 0.096 \mu\text{M}$ ) compared to the shorter H3(1-15)K4 peptide bearing only unmodified K4 ( $K_D = 8.6 \mu\text{M}$ ) or for the H3(13-32)K23ac peptide



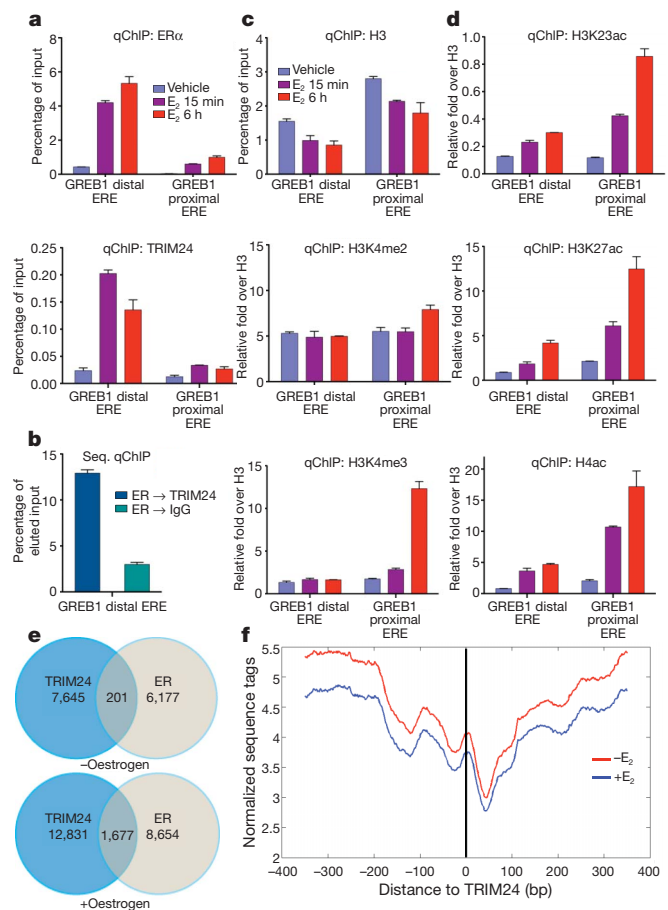
bearing only acetylated K23 marks ( $K_D = 8.8 \mu\text{M}$ ). Without acetylation on K23, the binding for H3(1–33)K4 is 24-fold weaker (Fig. 2e;  $K_D = 2.3 \mu\text{M}$ ); when K4 is tri-methylated, the binding for H3(1–33)K4me3K23ac is sixfold weaker (Fig. 2e;  $K_D = 0.56 \mu\text{M}$ ). Similarly, mutants that disrupt either the PHD finger binding pocket (C840W) or the bromodomain binding pocket (F979A/N980A) also decreased binding for H3(1–33)K4K23ac peptide by 6–7 fold (Fig. 2e and Supplementary Table 2).

On the basis of fluorescence polarization (FP)-based measurement, wild-type TRIM24 PHD-Bromo also showed strong binding affinity for H3(1–33)K4K23ac peptide ( $K_D = 0.185 \mu\text{M}$ ); peptides trimethylated at K4 or without acetylation at K23 displayed 13–23 fold weaker interaction (Fig. 2f and Supplementary Table 3). Mutation on the PHD finger binding pocket (C840W) or the bromodomain binding pocket (F979A/N980A) showed similar decrease in binding affinities (Fig. 2f and Supplementary Table 3). These binding data strongly support our structural results, which indicate that unmodified H3K4 and acetylated H3K23 are a pair of natural histone marks targeted by TRIM24 PHD-Bromo that can be read in a combinatorial manner on a single histone peptide. This combinatorial readout can greatly increase the recruitment of TRIM24 to nucleosomes bearing these two marks.

### TRIM24 and ER $\alpha$ recruitment to chromatin

Combinatorial histone modifications of unmethylated H3K4 alongside acetylated lysines have no straightforward interpretation by the models of chromatin modification and regulated activation or repression of transcription. We considered a model where TRIM24 regulates gene expression by specific binding to chromatin with non-canonical combinations of PTMs, and focused on co-regulation of ER $\alpha$ . This was because *in vitro* interactions between TRIM24 and nuclear receptors, including ER $\alpha$ , are ligand-dependent (Supplementary Fig. 10 and ref. 26), and because ligand-activated, ER-response elements (EREs) are notably independent of H3K4me2 and H3K4me3 modifications<sup>27,28</sup>. We used chromatin immunoprecipitation (ChIP) and sequential ChIP analyses of ER $\alpha$ -positive, MCF7 breast cancer cells to assess whether TRIM24 is recruited with ER $\alpha$  to specific EREs of the *GREB1*, *PR* and *pS2/TFF1* genes (Fig. 3a, b, and Supplementary Fig. 11). Oestrogen-activated recruitment occurs robustly within 15 min, and by six hours yields a sevenfold increase of ER $\alpha$  binding and a sixfold increase of TRIM24 binding at the *GREB1* distal ERE, ~40 kilobases (kb) upstream of the transcription start site (Fig. 3a). ChIP analysis of H3K4me2/3 after oestrogen treatment indicates that quantified H3K4me2 and H3K4me3 levels decreased at distal ERE sites (Supplementary Fig. 12 and ref. 27) and, when normalized for nucleosomal occupancy, decreased or was unchanged at distal EREs (Fig. 3c and Supplementary Fig. 13). Importantly, TRIM24 is recruited in the absence of changes in H3K4 methylation. In contrast, H3K23ac, H3K27ac and H4ac, which are targeted by the TRIM24 bromodomain, are enriched at both distal and proximal EREs after oestrogen-addition (Fig. 3d). These findings suggest that TRIM24 interacts with ER $\alpha$  and chromatin lacking H3K4 methylation but enriched in lysine acetylation, as suggested by our structural analyses, in response to oestrogen.

These findings stand in contrast to a model of chromatin accessibility at ER binding sites, facilitated by FOXA1 and H3K4me2 enrichment in response to oestrogen treatment<sup>29</sup>, but are in agreement with findings that H3K4me3 is not present at a majority of distal ERE regions<sup>28</sup>. We evaluated global chromatin-association of TRIM24, ER $\alpha$  and H3K4me2, by ChIP and deep sequencing of antibody-enriched DNA fragments (ChIP-seq). These analyses revealed binding of TRIM24 and ER at more than 10,000 sites genome-wide; half of which, in each case, are oestrogen-dependent (Fig. 3e and Supplementary Fig. 14a). Shared target sites of ER $\alpha$  and co-regulator TRIM24 increase dramatically (eightfold) in response to oestrogen (Supplementary Fig. 14b), are highly enriched ( $P$  value  $< 0.001$ ) at genes regulated by oestrogen<sup>30</sup> (Supplementary Fig. 14c), and function in cell cycle, kinase activity



**Figure 3 | TRIM24 is recruited with ER $\alpha$  to ERE sites depleted of H3K4me2.**

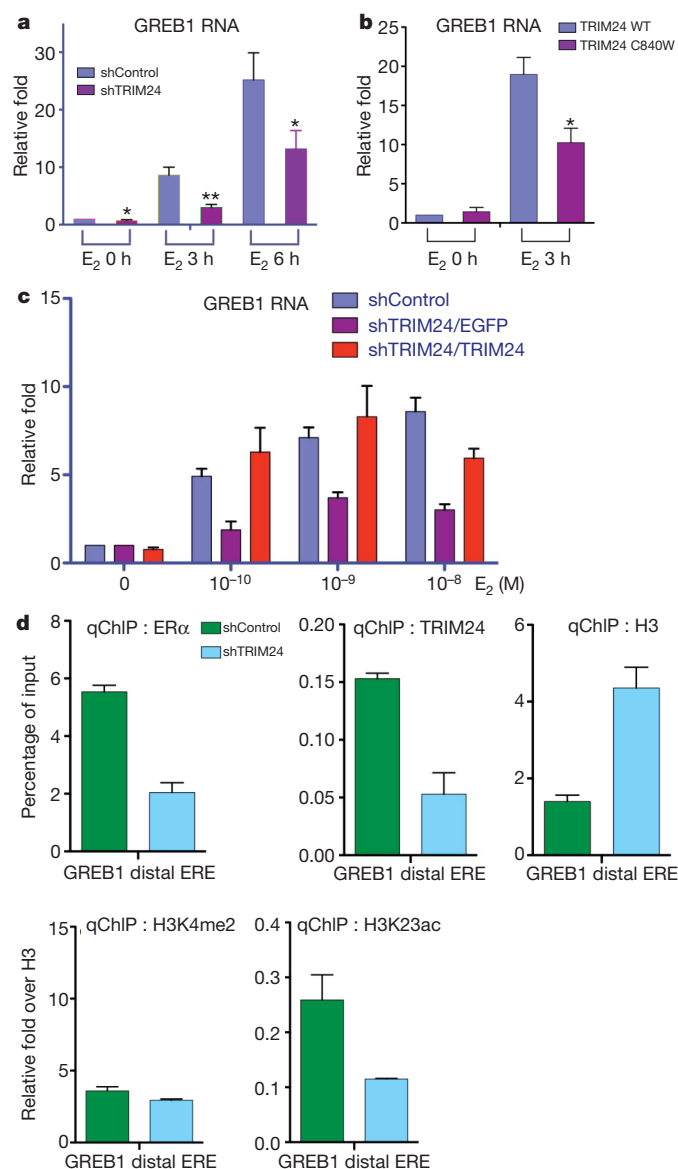
**a**, ChIP of ER $\alpha$  and TRIM24 binding at EREs of *GREB1*, 15 min and 6 h after treatment with oestradiol ( $E_2$ ). Vehicle, EtOH. **b**, Sequential-ChIP: ER $\alpha$  and TRIM24, 6 h after  $E_2$  addition. **c, d**, ChIP for H3 and histone modifications, 15 min and 6 h after  $E_2$ , normalized for H3. Each bar represents averaged results,  $n = 3$  biological replicates, assayed 3 times each; error bars, s.d. **e**, Genome-wide TRIM24 and ER $\alpha$  binding sites in MCF7 cells,  $-E_2$  or  $+E_2$ . Two independent experiments analysed. **f**, Normalized genome-wide H3K4me2 within a window of 800 bp, centred at TRIM24 binding sites (designated as 0),  $+E_2$  (blue line) or  $-E_2$  (red line).

and signal transduction (DAVID analyses<sup>31</sup>, Supplementary Table 4). Biological pathway analysis (Ingenuity Systems, [www.ingenuity.com](http://www.ingenuity.com)) revealed that multiple gene targets of TRIM24 are associated with breast cancer (Supplementary Tables 5 and 6). The number of target sites shared by TRIM24 and ER $\alpha$  (1,677 sites) is similar to ER $\alpha$  and FOXA1 (ref. 29), with little overlap among all three (263 sites) (Supplementary Fig. 14b). Consistent with our structural analyses, TRIM24 binding occurs globally at sites depleted of H3K4me2 (Fig. 3f and Supplementary Figs 14d and 15). Thus, ER $\alpha$ -regulated genes may be divided into multiple classes, defined by specific co-regulators and their dependence on H3K4 methylation.

### TRIM24 is overexpressed in breast cancer

Depletion of TRIM24 caused a significant decrease in ER $\alpha$ -mediated activation of *GREB1*, *PR* and *pS2* gene expression (Fig. 4a and Supplementary Fig. 16a). Importantly, re-introduction of wild type (WT), but not PHD finger mutant (C840W), TRIM24 fully restored ER $\alpha$ -mediated transcription activation (Fig. 4b), and enabled ER $\alpha$ -response at lower levels of hormone (Fig. 4c). Decreased ER $\alpha$ -mediated activation is due to loss of TRIM24-dependent ER $\alpha$ -interactions with chromatin (Fig. 4d and Supplementary Fig. 17), without alternation of ER $\alpha$  expression (Supplementary Fig. 16b). H3K4me2/3 levels at the distal ERE of *GREB1* lack hormone responsiveness and are TRIM24-independent

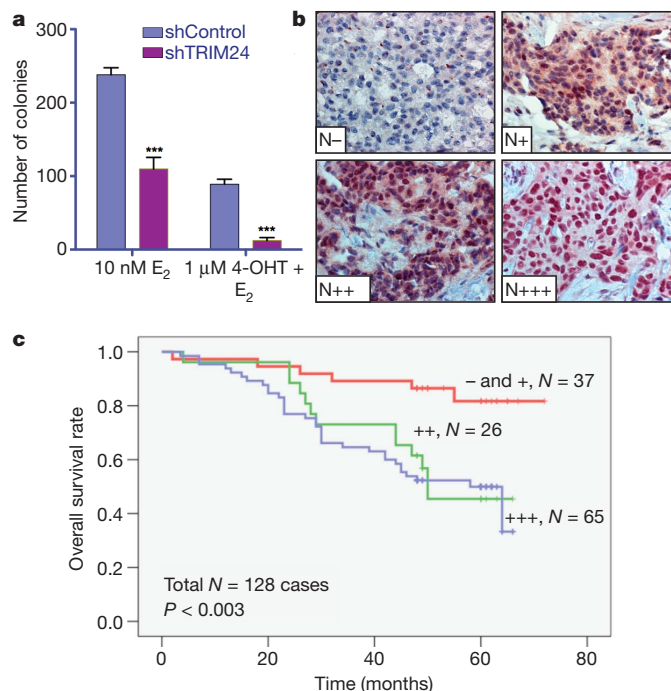




**Figure 4 | TRIM24 functions as a co-activator and stabilizes ER $\alpha$ -chromatin interactions.** **a**, Stable shControl and shTRIM24 MCF7 cells  $\pm$  E<sub>2</sub>. \* $P < 0.05$ ; \*\* $P < 0.01$ . **b**, TRIM24 WT and TRIM24 C840W expressed in stable shTRIM24 MCF7 cells  $\pm$  E<sub>2</sub>. **c**, shControl and shTRIM24 MCF7 cells, E<sub>2</sub> range. TRIM24 WT or EGFP control expressed in shTRIM24 MCF7 cells. In **a**, **b** and **c**, GREB1 RNA levels are normalized to GAPDH; untreated shControl MCF7 is set as one. Each bar is an average of 3 biological replicates, 3 independent RT-PCR assays of each; error bars, s.d. **d**, ChIP of ER $\alpha$  and TRIM24 binding, histone H3 and histone modifications, 6 h after E<sub>2</sub> addition, shControl and shTRIM24 MCF7 cells. Histone modifications normalized for H3 recovery. Each bar represents averaged results,  $n = 3$  and 3 assays of each; error bars, s.d.

(Fig. 4d and Supplementary Fig. 16c). In contrast, nucleosomal occupancy at EREs is increased alongside decreased acetylation of H4, H3K23 and H3K27, reflecting loss of ER $\alpha$ -activated chromatin structure (Fig. 4d and Supplementary Fig. 16c).

Depletion of TRIM24 led to reduced survival and proliferation of tumour-derived breast cancer cells, and is highly additive with 4-OH-tamoxifen, an inhibitor of ER $\alpha$ <sup>32</sup> (Fig. 5a). We immunostained tissue samples from a breast cancer patient cohort to assess the impact of TRIM24 expression on breast cancer survival (Fig. 5b). In 128 cases of non-metastatic breast cancer, expression of TRIM24 fell into four classes: N– and N+, undetectable to low level in few foci (29%); N++, abundant foci with expression in nuclear and cytoplasmic



**Figure 5 | Aberrant expression of TRIM24 correlates with poor survival of breast cancer patients.** **a**, shControl and shTRIM24 MCF7 cells with E<sub>2</sub> or E<sub>2</sub> + 4-OH-tamoxifen (4-OHT), as indicated. Each bar represents the averaged results for three independent colony formation assays in triplicate plates; error bars, s.d. \* $P < 0.0001$ . **b**, Immunohistochemistry: 128 surgical specimens of breast cancer immunostained for TRIM24: subcellular localization (N) and staining intensity (strong, +++; moderate, ++; weak or slightly above background, +; none, –). **c**, The overall survival rate of 128 patients with non-metastatic disease, classified by TRIM24 expression (as in **b**), plotted by the Kaplan–Meier method.

compartments (20%); and N+++, abundant foci with high expression in nuclei (51%). Overexpression of TRIM24 (+++, ++) is clearly correlated with poor patient survival, independent of ER status (Fig. 5c and Supplementary Table 7).

## Discussion

Our identification of the PHD-Bromo as a reader of H3K4me0 and H3K23ac within a single histone tail, or of H3K4me0 and non-contiguous acetylated lysines, suggests that TRIM24 may have multiple roles in chromatin regulation<sup>20</sup>. TRIM24 is a co-activator of ER $\alpha$  at distal EREs, a platform well suited for stable interactions with TRIM24 PHD-Bromo. ER $\alpha$  recruits histone acetyltransferases—for example, CBP/p300, GCN5 and P/CAF (ref. 33)—to acetylate histones. LSD1 (KDM1), a biochemically and structurally characterized demethylase for H3K4me2/1 (refs 34, 35) and an androgen-regulated demethylase of H3K9me (ref. 36), is resident<sup>37</sup> or rapidly recruited<sup>27</sup> to EREs where H3K4 remains depleted of methylation even with oestrogen activation (Fig. 3c, Supplementary Fig. 18 and ref. 28). These parallel processes establish a combinatorial histone signature with high affinity for TRIM24 binding to chromatin.

Aberrant expression of TRIM24 may promote tumour development and progression by multiple mechanisms of dysfunction. TRIM24 is a potent co-activator of ER $\alpha$ , which is associated with cellular proliferation and neoplasia in breast cells<sup>38,39</sup>, and a negative regulator of p53 stability<sup>19</sup>. TRIM24 is a target of chromosomal translocations to form oncogenic fusion proteins in acute promyelocytic leukaemia<sup>40</sup>, papillary thyroid carcinoma<sup>41</sup> and myeloproliferative syndrome<sup>42</sup>. Here we have shown that TRIM24 expression is directly correlated with poor patient survival in both ER-positive and ER-negative breast cancer. These results suggest that TRIM24 is a dual

domain histone reader with considerable potential as a therapeutic target in multiple cancers.

## METHODS SUMMARY

Wild-type and mutant forms of TRIM24 PHD-Bromo were expressed in *Escherichia coli* and purified to homogeneity. Histone biotinylated peptides or purified histone proteins were incubated with GST-proteins, and bound proteins detected by immunoblotting. All crystals were obtained by the hanging-drop method at 20 °C; structures were solved by the molecular replacement method, and refined with cycled model building and refinement procedures. Histone peptides with or without biotin labelling were used for ITC binding. Fluorescein-labelled peptides were used for fluorescence polarization analysis. Stable short hairpin (sh)Control and shTRIM24 MCF7 cells were maintained with 2.5 µg ml<sup>-1</sup> puromycin and, for hormone treatment, were grown in hormone-free media for 96 h before addition of ethanol or 10 nM oestradiol (Sigma) for indicated times. Global expression analyses and calculation of enrichment of shared TRIM24 and ERα binding at oestrogen-regulated genes<sup>30</sup> were determined, and validated by real-time RT-PCR. Surgical specimens of breast cancer from 128 non-metastatic patients were immunostained for TRIM 24 (TRIM24 antibody, Proteintech Group), and scored by subcellular localization, staining intensity, and fraction of positive staining. The overall survival after surgery was plotted by the Kaplan–Meier method<sup>43</sup>.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 29 January; accepted 30 September 2010.

1. Taverna, S. D. *et al.* How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nature Struct. Mol. Biol.* **14**, 1025–1040 (2007).
2. Jenuwein, T. & Allis, C. D. Translating the histone code. *Science* **293**, 1074–1080 (2001).
3. Ruthenburg, A. J., Li, H., Patel, D. J. & Allis, C. D. Multivalent engagement of chromatin modifications by linked binding modules. *Nature Rev. Mol. Cell Biol.* **8**, 983–994 (2007).
4. Berger, S. L. The complex language of chromatin regulation during transcription. *Nature* **447**, 407–412 (2007).
5. Baker, L. A., Allis, C. D. & Wang, G. G. PHD fingers in human diseases: disorders arising from misinterpreting epigenetic marks. *Mutat. Res.* **647**, 3–12 (2008).
6. Jacobson, R. H., Ladurner, A. G., King, D. S. & Tjian, R. Structure and function of a human TAF<sub>II</sub>250 double bromodomain module. *Science* **288**, 1422–1425 (2000).
7. Morinier, J. *et al.* Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. *Nature* **461**, 664–668 (2009).
8. Lan, F. *et al.* Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated gene repression. *Nature* **448**, 718–722 (2007).
9. Org, T. *et al.* The autoimmune regulator PHD finger binds to non-methylated histone H3K4 to activate gene expression. *EMBO Rep.* **9**, 370–376 (2008).
10. Ragvin, A. *et al.* Nucleosome binding by the bromodomain and PHD finger of the transcriptional cofactor p300. *J. Mol. Biol.* **337**, 773–788 (2004).
11. Zhou, Y. & Grummt, I. The PHD finger/bromodomain of NoRC interacts with acetylated histone H4K16 and is sufficient for rDNA silencing. *Curr. Biol.* **15**, 1434–1438 (2005).
12. Peña, P. V. *et al.* Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature* **442**, 100–103 (2006).
13. Li, H. *et al.* Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* **442**, 91–95 (2006).
14. Lange, M. *et al.* Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex. *Genes Dev.* **22**, 2370–2384 (2008).
15. Koh, A. S. *et al.* Aire employs a histone-binding module to mediate immunological tolerance, linking chromatin regulation with organ-specific autoimmunity. *Proc. Natl Acad. Sci. USA* **105**, 15878–15883 (2008).
16. Hung, T. *et al.* ING4 mediates crosstalk between histone H3 K4 trimethylation and H3 acetylation to attenuate cellular transformation. *Mol. Cell* **33**, 248–256 (2009).
17. Wang, Z. *et al.* Proisomerization in MLL1 PHD3-bromo cassette connects H3K4me readout to CyP33 and HDAC-mediated repression. *Cell* **141**, 1183–1194 (2010).
18. Wang, G. G. *et al.* Haematopoietic malignancies caused by dysregulation of a chromatin-binding PHD finger. *Nature* **459**, 847–851 (2009).
19. Allton, K. *et al.* Trim24 targets endogenous p53 for degradation. *Proc. Natl Acad. Sci. USA* **106**, 11612–11616 (2009).
20. Poleshko, A. *et al.* Identification of a functional network of human epigenetic silencing factors. *J. Biol. Chem.* **285**, 422–433 (2009).
21. Meroni, G. & Diez-Roux, G. TRIM/RBCC, a novel class of 'single protein RING finger' E3 ubiquitin ligases. *Bioessays* **27**, 1147–1157 (2005).
22. Raymond, A. *et al.* The tripartite motif family identifies cell compartments. *EMBO J.* **20**, 2140–2151 (2001).
23. LeDouarin, B. *et al.* A possible involvement of TIF1 alpha and TIF1 beta in the epigenetic control of transcription by nuclear receptors. *EMBO J.* **15**, 6701–6715 (1996).
24. Soliman, M. A. & Riabowol, K. After a decade of study-ING, a PHD for a versatile family of proteins. *Trends Biochem. Sci.* **32**, 509–519 (2007).
25. Peña, P. V. *et al.* Histone H3K4me3 binding is required for the DNA repair and apoptotic activities of ING1 tumor suppressor. *J. Mol. Biol.* **380**, 303–312 (2008).
26. Thénot, S., Henriquet, C., Rochefort, H. & Cavaillès, V. Differential interaction of nuclear receptors with the putative human transcriptional coactivator hTIF1. *J. Biol. Chem.* **272**, 12062–12068 (1997).
27. Garcia-Bassets, I. *et al.* Histone methylation-dependent mechanisms impose ligand dependency for gene activation by nuclear receptors. *Cell* **128**, 505–518 (2007).
28. Fullwood, M. J. *et al.* An oestrogen-receptor-α-bound human chromatin interactome. *Nature* **462**, 58–64 (2009).
29. Lupien, M. *et al.* FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. *Cell* **132**, 958–970 (2008).
30. Lin, C. Y. *et al.* Whole-genome cartography of estrogen receptor α binding sites. *PLoS Genet.* **3**, e87 (2007).
31. Huang, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols* **4**, 44–57 (2009).
32. Eckert, R. L. & Katzenellenbogen, B. S. Physical properties of estrogen receptor complexes in MCF-7 human breast cancer cells. Differences with anti-estrogen and estrogen. *J. Biol. Chem.* **257**, 8840–8846 (1982).
33. Yanagisawa, J. *et al.* Nuclear receptor function requires a TIFC-type histone acetyltransferase complex. *Mol. Cell* **9**, 553–562 (2002).
34. Stavropoulos, P., Blobel, G. & Hoelz, A. Crystal structure and mechanism of human lysine-specific demethylase-1. *Nature Struct. Mol. Biol.* **13**, 626–632 (2006).
35. Shi, Y. *et al.* Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**, 941–953 (2004).
36. Wissmann, M. *et al.* Cooperative demethylation by JMJD2C and LSD1 promotes androgen receptor-dependent gene expression. *Nature Cell Biol.* **9**, 347–353 (2007).
37. Perillo, B. *et al.* DNA oxidation as triggered by H3K9me2 demethylation drives estrogen-induced gene expression. *Science* **319**, 202–206 (2008).
38. Katzenellenbogen, B. S. Estrogen receptors: bioactivities and interactions with cell signaling pathways. *Biol. Reprod.* **54**, 287–293 (1996).
39. Cheskis, B. J., Greger, J. G., Nagpal, S. & Freedman, L. P. Signaling by estrogens. *J. Cell. Physiol.* **213**, 610–617 (2007).
40. Zhong, S. *et al.* A RA-dependent, tumour-growth suppressive transcription complex is the target of the PML-RARα and T18 oncoproteins. *Nature Genet.* **23**, 287–295 (1999).
41. Klugbauer, S. & Rabes, H. M. The transcription coactivator HTIF1 and a related protein are fused to the RET receptor tyrosine kinase in childhood papillary thyroid carcinomas. *Oncogene* **18**, 4388–4393 (1999).
42. Belloni, E. *et al.* 8p11 myeloproliferative syndrome with a novel t(7;8) translocation leading to fusion of the FGFR1 and TIF1 genes. *Genes Chromosom. Cancer* **42**, 320–325 (2005).
43. Xia, W. *et al.* Phosphorylation/cytoplasmic localization of p21Cip1/WAF1 is associated with HER2/neu overexpression and provides a novel combination predictor for poor prognosis in breast cancer patients. *Clin. Cancer Res.* **10**, 3815–3824 (2004).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This work was supported by funds from the National Institutes of Health (NIH GM081627) and the George and Cynthia Mitchell Foundation (to M.C.B.), from NIH (U54 RR025216 and P30DK078392-01) to B.A., from NIH (GM079641) to O.G., from the Sister Institution Fund of China Medical University and Hospital and MDACC to M.-C.H., from the Starr Foundation and the Leukemia and Lymphoma Society to D.J.P., from the Max Planck Society to W.F., and from the NCI Cancer Center (Support Grant) to the UT MD Anderson Cancer Center. W.-W.T. was supported in part by the Sowell-Huggins Foundation; S.W. by a long-term EMBO fellowship; T.T.Y. by T32 HD07325; and K.C.A. by the Center for Cancer Epigenetics. We thank J. Song, D.C. Jamison, A. Dose, Z. Coban and Y. Wei for technical support and assistance. We are grateful to S. Stratton, M. Lee, M. Bedford, G. Lozano, S. Dent, A. Nardulli and members of our laboratories for advice, reagents and discussions.

**Author Contributions** W.-W.T. identified ER interactions, and performed molecular biology and IHC studies; Z.W. solved the molecular structures of TRIM24 PHD-Bromo in the free and bound states, and performed ITC binding affinity studies; T.T.Y. performed mutagenesis, ChIP and clonogenic analyses; C.-Y.T. performed clonogenic assays; K.C.A. performed bioinformatic analyses; W.X. analysed patient samples; X.S. performed peptide array analyses; S.W., D.S. and W.F. performed and analysed FP experiments; O.G., B.A., W.P., W.F., M.-C. H., D.J.P. and M.C.B. discussed studies; and D.J.P. and M.C.B. designed structural and functional studies, analysed data and wrote the paper. W.-W.T. and Z.W. contributed equally to this work. All authors discussed and commented on the manuscript.

**Author Information** The X-ray coordinates of TRIM24 PHD-Bromo in the free state and when bound to H3(1–10)K4, H3(13–32)K23ac, H3(23–31)K27ac and H4(14–19)K16ac peptides have been deposited in the Protein Data Bank (PDB) under accession numbers 3033, 3037, 3034, 3035 and 3036, respectively. ChIP-sequencing files and data are deposited at the NCBI Gene Expression Omnibus (GEO) site as accession number GSE24166. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to D.J.P. (pateld@mskcc.org) or M.C.B. (mbarton@mdanderson.org).

## METHODS

**Cell culture.** MCF7 and HEK293T cells were obtained from ATCC. Stable shControl and shTRIM24 MCF7 cells were maintained with  $2.5 \mu\text{g ml}^{-1}$  puromycin. For hormone treatment, MCF7 cells were plated under normal growth condition (DMEM with 10% FBS and  $1 \times$  penicillin/streptomycin). When cell density reached 25% confluence, the media were changed to hormone-free media (phenol red-free DMEM, 10% charcoal dextran-treated FBS) for 96 h. The hormone-deprived MCF7 cells were treated with ethanol or 10 nM oestradiol (Sigma) for indicated times. The TRIM24 shRNA for stable cell line creation was AAGCA GGTGG AACAG GATAT TAAAG TTGC.

**Biotinylated peptide pulldown assay.** Histone biotinylated peptides (1  $\mu\text{g}$ , Millipore) were incubated with GST-proteins (2  $\mu\text{g}$ ), in 500  $\mu\text{l}$  NTP buffer (300 mM NaCl, 50 mM Tris-HCl, 0.1% NP-40, pH 7.4) overnight at  $4^\circ\text{C}$ . 20  $\mu\text{l}$  of a 50% slurry of Streptavidin-coated beads (Thermo) were added and incubated for 1 h at  $4^\circ\text{C}$ . The beads were recovered by centrifugation and washed 6 times (10 min at  $4^\circ\text{C}$ ) with NTP. The peptide/protein bound beads were resuspended in  $3 \times$  SDS-PAGE loading buffer, heated for 5 min at  $95^\circ\text{C}$  and separated on 4–12% gradient polyacrylamide gels (Invitrogen). GST-proteins were detected by western blotting analysis with anti-GST antibody (Cell Signaling, dilution 1:1,000).

**GST protein pull-down assay, co-IP and transient transfection.** GST-proteins (2  $\mu\text{g}$ ) were incubated with purified histone proteins (10  $\mu\text{g}$ ) in 500  $\mu\text{l}$  NTP overnight at  $4^\circ\text{C}$ . 25  $\mu\text{l}$  of a 50% slurry of GST-beads were added and incubated for 2 h at  $4^\circ\text{C}$ , recovered by centrifugation and washed 4 times (10 min at  $4^\circ\text{C}$ ) with NTP. The protein bound beads were separated by SDS-PAGE and subjected to western blotting analysis. Anti-GST (Cell Signaling), anti-TRIM24 (Novus Biological), anti-histone H3 (Abcam), anti-H3K9me2 (Active Motif) and anti-H3K4me3 (Active Motif) antibodies were used in immunoblotting.

Ethidium bromide-treated and precleared cell lysate was used for IP with TRIM24 antibody (2.5  $\mu\text{l}$ , Novus Biological) or control rabbit IgG antibody (2.5  $\mu\text{l}$ , Upstate) as previously described<sup>44</sup>. Then anti-ER (F-10, Santa Cruz) and anti-TRIM24 were used in western blotting analysis. Transient transfections were performed as previously described<sup>45</sup>, but with Effectene (Qiagen) used according to manufacturer's directions.

**Chromatin immunoprecipitation assay (ChIP).** ChIP assays were performed as described<sup>45</sup>, with minimal modification. Briefly, treated MCF7 cells were cross-linked with 1% formaldehyde for 15 min. The cross-linking reaction was stopped with 0.125 M glycine. The cross-linked cells were washed with PBS three times and stored at  $-80^\circ\text{C}$  before use. The fragmented, precleared chromatin lysate was incubated overnight with specific antibodies: ER (F-10, Santa Cruz), TRIM24 (Novus Biological), histone H3 (Abcam), H3K4me2 (Active Motif), H3K4me3 (Active Motif), H3K23ac (Active Motif), H3K27ac (Active Motif), H4ac (Upstate/Millipore) or normal sheep IgG (Upstate/Millipore). To analyse specific, antibody- and protein-bound DNA, qPCR was conducted in a 7500 FAST ABI instrument. PCR primers were used as described<sup>27,46</sup>.

**ChIP-sequencing analysis.** Purified ChIP DNA was prepared for sequencing using the Illumina ChIP-Seq sample preparation kit (IP-102-1001, Illumina). Sequencing was performed on an Illumina Genome Analyser II using 36 cycles. Sequences were aligned to human genome release hg18 (ref. 47) using the ELAND software at the default 0 to 2 allowed mismatch setting<sup>48</sup>.

High-quality aligned read data was converted to the BED format and analysed for peaks using the MACS software programs<sup>49</sup>. Peaks showing altered pull-down patterns were mapped to the genome, and a list of candidate genes flanking each peak was generated. Foxa1 (ChIP-Chip) binding regions<sup>29</sup> were obtained from <http://research.dfci.harvard.edu/brownlab/datasets/>. Overlapping regions were identified if any of two regions share more than 80% of their length with the other region. Venn diagrams were prepared on the basis of that assumption.

To identify possible target genes, a RefSeq data file was downloaded from UCSC genome browser (hg18). Genes within 10 kb of the binding regions were marked as target genes; DAVID<sup>51</sup> was used to assess biological functions of targets, and Ingenuity Pathway Analysis (IPA) software was used to ascertain the pathways that are enriched by these target genes (Ingenuity Systems, [www.ingenuity.com](http://www.ingenuity.com)). For H3K4me2 signals around the Trim24 binding regions, in the presence or absence of oestrogen, numbers of fragments that overlap each position were counted. The Input signal was calculated by the same procedure and overall H3K4me2 signal values were normalized by Input signal.

For intersection with global expression analyses and calculation of enrichment of shared Trim24 and ER binding at oestrogen-regulated genes<sup>30</sup>, raw CEL files were downloaded from the GEO database. RMA normalization was used with default options (with background correction, quantile normalization, and log transformation) to normalize the intensity of each probeset. The SAM statistical method was used to select differentially expressed genes, based on a  $q$ -value of less than 2%. After filtration, a total of 1,887 genes were selected as differentially expressed after oestrogen treatment. Venn diagrams were formed based on the

proximity of binding regions to genes within 10 kb of the binding sites.  $P$ -values were calculated based on binomial tests.

**Real time RT-PCR analysis.** RNA was isolated at indicated time points using Trizol reagent (Invitrogen) following manufacturer's instructions. cDNA was synthesized using a RT-PCR kit (Invitrogen) according to manufacturer's instructions. Synthesized cDNA was treated with RNaseH before performing PCR reactions. Real time PCR reactions were performed in an Applied Biosystems 7500 Fast Real time PCR instrument. Generally, 10  $\mu\text{l}$  PCR reactions were set up in 96-well plates. The reaction mixtures contain 2  $\mu\text{l}$  diluted cDNA (1:10 dilution), 0.25  $\mu\text{l}$  each of forward and reverse primers (20  $\mu\text{M}$ ), 5  $\mu\text{l}$   $2 \times$  SYBR Green Mix and 2.5  $\mu\text{l}$  water. PCR primers were used as described<sup>46</sup>.

GraphPad Prism5 software was used for analysis of  $P$ -values based on at least two independent experiments in three independent PCR reactions. The two-tailed paired  $t$ -test was used to compare the differences between two groups for relative fold. The two-tailed unpaired  $t$ -test was used to compare the differences between two groups for actual percentage.  $P$ -values  $<0.05$  were considered statistically significant.

**Colony formation assay.** 500 stable control shRNA depleted or TRIM24 shRNA depleted MCF7 cells were distributed in DMEM with 10% FBS and  $1 \times$  P/S and then placed in 60 mm<sup>2</sup> plates at  $37^\circ\text{C}$  for 14 days. Colonies were stained by crystal violet 14 days after seeding. Colonies of  $>50$  cells were counted using a dissecting microscope. Each sample was performed in triplicate.

**Clinical patient samples and immunohistochemistry.** We obtained 134 archived blocks containing formalin-fixed, paraffin-embedded infiltrating breast carcinoma from the Department of Pathology, Shanghai East Breast Disease Hospital, China. All of the patients were women with non-metastatic disease who had undergone mastectomy and axillary lymph node dissection between 1988 and 1994. After surgical treatment, the patients were offered adjuvant chemotherapy and/or radiotherapy and hormone therapy, depending on the number of lymph node metastases, status of menopause, and oestrogen and/or progesterone receptor positivity. The clinicopathological characteristics of the study population, including age, tumour size, lymph node status, tumour grade, and oestrogen receptor/progesterone receptor positivity, were obtained from medical records. The stage was assessed by the TNM clinical staging system of the American Joint Committee on Cancer. Patients were followed for 4–72 months, with a median follow-up of 48 months. We have follow-up information from 128 patients. Most of the tissues in this cohort were described previously<sup>43</sup>, and used as approved by the Institutional Review Board.

128 surgical specimens of breast cancer were immunostained for TRIM 24 (TRIM24 antibody, Proteintech Group). Immunoreactivity of the TRIM24 antibody was scored according to the subcellular localization (nuclear and/or cytoplasmic), staining intensity (strong + + +, moderate + +, weak + and faint or slightly above background), and fraction of positive staining. The mean fraction of positive tumour cells was determined in at least nine areas at  $\times 100$  or  $\times 200$  magnification. Immunoreactivity scores were as described previously. For example, 0 is no staining; + is weak staining; ++ is moderate staining; +++ is strong staining and the values for percentage of positive tumour cells: 0 = 0–0.9%; 1 = 1–10%; 2 = 11–50%; 3 = 51–80%; 4 = 81–100%. Magnification was  $\times 400$ ; at least 200 tumour cells were counted<sup>43</sup>. ER assessment was done in a similar manner.

Data on eligible patients were summarized by use of standard descriptive statistics and frequency tabulation. The overall survival after surgery was plotted by use of the Kaplan–Meier method. The log-rank test was used to analyse differences in survival time. All tests were two-sided and the level of significance was set at 0.05.

**Histone peptide microarray hybridization.** Peptide microarray assays were performed as described<sup>50</sup>. Briefly, biotinylated histone peptides were printed in hexaplicates onto a streptavidin coated slide (ArrayIt) using a VersArray Compact Microarrayer (BioRad). After short blocking with biotin (Sigma), the slides were incubated with the GST-TRIM24 PHD-Bromo in binding buffer (300 mM NaCl, 50 mM Tris-HCl 7.5, 0.1% NP-40, 1 mM PMSF, 20% fetal bovine serum) overnight at  $4^\circ\text{C}$  with gentle agitation. After washing with the same buffer, slides were probed with anti-GST antibody and then fluorescein-conjugated secondary antibody and visualized with a GenePix 4000 scanner (Molecular Devices).

**Protein preparation.** As the long linker (30 residues) between TRIM24 PHD finger and bromodomain was predicted to be a loop, we first expressed PHD finger (824–886) and bromodomain (896–1016) separately, and then checked their interaction by the gel-filtration method. The mixture of PHD finger and bromodomain contained a higher molecular weight peak that corresponded to the complex fraction (Supplementary Fig. 16), which indicates that TRIM24 PHD-Bromo forms a structural unit with direct interactions between the domains. Next, we focused on the expression and crystallization of the TRIM24 PHD-Bromo dual domain. We cloned different lengths of TRIM24



PHD-Bromo fragments into pRSFDuet-1 vector (modified pRSFDuet-1 vector with 6×His plus yeast sumo as fusion tag). Fragment 824–1011 was used for all the binding assays. Further, all the mutations were introduced into the same construct using a QuikChange Kit (Stratagene). Fragment 824–1006 was used for all crystallographic studies. Protein expression was carried out in BL21 (DE3) *E. coli* cells. Cells were grown until A<sub>600</sub> reached around 1.0, then the media were cooled and 0.2 mM IPTG and 0.1 mM ZnCl<sub>2</sub> were added to the culture to induce the protein expression at 25 °C for 8 h.

*E. coli*-expressed wild-type and mutant TRIM24 PHD-Bromo were affinity purified by nickel-charged HisTrap Chelating FF columns (GE Healthcare). *E. coli* cells were disrupted by sonication in loading buffer (20 mM Tris (pH 8.0), 500 mM NaCl, 20 mM imidazole). After centrifugation, the supernatant was loaded onto a nickel column and was washed extensively with loading buffer. The target protein was eluted with a linear gradient of 50–500 mM imidazole. The target protein was collected and dialysed overnight in loading buffer, with Ulp1 protease added to cleave the 6×His-sumo tag. Digested protein was loaded onto a nickel-column again to remove the 6×His-sumo tag and the His-tagged protease. The flow through containing TRIM24 PHD-Bromo was collected and further purified by a Hiload 75 26/60 column in buffer containing 10 mM Tris (pH 8.0), 100 mM NaCl and 5 mM DTT. The major peak was pooled, concentrated to high concentration and stored in a –80 °C freezer.

**Crystallization.** Crystals of TRIM24 PHD-Bromo in the free form were grown by mixing equal volumes of 20 mg ml<sup>−1</sup> protein and crystallization buffer (100 mM Hepes (pH 7.5), 2.0 M ammonium sulphate and 2% polyethylene glycol 400) at 20 °C. Crystals appeared after two days and reached their full lengths within a week.

Crystals of the complex of TRIM24 PHD-Bromo and unmodified H3(1–10)K4 peptide were grown by incubating 20 mg ml<sup>−1</sup> protein with peptide at a 1:1.5 ratio, followed by mixing with an equal volume of crystallization buffer (50 mM Tris (pH 7.5), 30% polyethylene glycol monomethyl ether 5000, and 100 mM NaCl) at 20 °C. Crystals appeared within a week and reached their full size in two weeks.

Crystals of the complex of TRIM24 PHD-Bromo and H3(13–32)K23ac peptide were grown by mixing 20 mg ml<sup>−1</sup> protein with peptide at 1:1.5 ratio, followed by mixing with an equal volume of crystallization buffer (100 mM sodium citrate (pH 5.6), 200 mM potassium/sodium tartrate tetrahydrate and 1.6 M ammonium sulphate) at 20 °C.

Crystals of the complex of TRIM24 PHD-Bromo and H3(23–31)K27ac peptide were grown by mixing 20 mg ml<sup>−1</sup> protein with peptide at a 1:3 ratio, followed by mixing with an equal volume of crystallization buffer (50 mM Tris (pH 8.0), 200 mM sodium acetate, 30% polyethylene glycol 4000) at 20 °C. Initially, we could only get twinned crystals that could not be used for structure determination. Later, single crystals were obtained by including detergent 0.8 mM CHAPSO (Hampton Research) into the crystallization buffer.

Crystals of the complex of TRIM24 PHD-Bromo and H4(14–19)K16ac peptide were prepared by mixing 20 mg ml<sup>−1</sup> protein with peptide at a 1:3 ratio, followed by mixing with an equal volume of crystallization buffer (50 mM Bis-Tris (pH 6.5), 30% polyethylene glycol 3350, 100 mM ammonium acetate) at 20 °C. Crystals appeared within a day and reached their full lengths in a week. Fresh crystals (grown within a week) were used for data collection, as crystals grown for longer than 3 weeks diffracted poorly. For all the crystals mentioned above, crystallization buffers with 20% glycerol were used as cryoprotectant.

**Data collection and structure determination.** Data sets for crystals of TRIM24 PHD-Bromo in the free state, as well as TRIM24 PHD-Bromo complexes with unmodified H3(1–10)K4, with H3(13–32)K23ac and with H4(14–19)K16ac peptides were collected at NE-CAT beamline 24ID-E, Advanced Photon Source, Chicago. Data sets for crystals of the complex of TRIM24 PHD-Bromo and H3(23–31)K27ac peptide were collected at beamline X29 at the National Synchrotron Light Source (NSLS, Brookhaven National Laboratories). All the data sets were integrated and scaled with the program HKL2000 suite. The free-form TRIM24 PHD-Bromo crystal belongs to *P*<sub>1</sub> space group and contains four molecules per asymmetric unit. The structure was solved by the molecular replacement method using the PHASER program<sup>51</sup>, using the BHC80 PHD finger (PDB coordinate 2PUY) and TRIM24 bromodomain (PDB coordinate 2YYN) as search models, while searching for four copies of each domain in the asymmetric unit. The initial model was rebuilt with COOT<sup>52</sup>, cycled with structure refinement by CNS<sup>53</sup>. The structures of all the complexes were solved by molecular replacement

with free-form TRIM24 PHD-Bromo as the model. All the crystallographic statistics are listed in Supplementary Table 1.

**Peptide synthesis and fluorescence polarization based measurement.** H3 peptides (residues 1–33) used for isothermal titration calorimetry (ITC) and for fluorescence polarization (FP) measurements were synthesized by Fmoc strategy using pseudo prolines at critical positions. Details of peptide synthesis are available on request. For FP assays, peptides were labelled using fluorescein-NHS ester (Invitrogen). Labelled peptides were purified on G10 gel filtration resin (GE Healthcare) and by RP-C18 HPLC. Single labelled species were identified by mass spectrometry. FP assays were essentially carried out and analysed as described using FP buffer (10 mM Tris-HCl, 100 mM NaCl, 5 mM DTT, pH 7.4)<sup>54</sup>. Titration series of 10 µl volume in 384-well plates were read multiple times on a Plate Chameleon II plate reader (HIDEX Oy). Multiple readings and independent titration series were averaged after data normalization.

**ITC measurements.** Calorimetric experiments were conducted at 25 °C with a MicroCal iTC200 instrument. Recombinant wild-type TRIM24 PHD-Bromo and its mutants were dialysed overnight against 20 mM Tris (pH 7.5), 50 mM NaCl and 2 mM β-mercaptoethanol. Aliquots of lyophilized peptides were dissolved in the same buffer before use. Calorimetric titration was performed by injecting synthetic peptide into wild-type TRIM24 PHD-Bromo or its mutants at various concentrations. Calorimetric titration data were fitted using Origin 7.0 software on the basis of a 1:1 binding stoichiometry.

**Peptides used for the crystallization and binding assays.** H3(1–10)K4, ARTKQTARKS; H3(1–21)K4, ARTKQTARKSTGGKAPRKQLAGGK-Biotin; H3(1–15)K4, ARTKQTARKSTGGKAY; H3(1–15)K4me1, ARTK(me1)QTARKSTGGKAY; H3(1–15)K4me2, ARTK(me2)QTARKSTGGKAY; H3(1–15)K4me3, ARTK(me3)QTARKSTGGKAY; H3(1–19)K14ac, ARTKQTARKSTGGK(ac)A PRKQ; H3(1–20)K9ac, ARTKQTARK(ac)STGGKAPRKQL; H3(23–31)K27ac, KAARK(ac)SAPA; H3(13–32)K27ac, GKAPRKQLATKAARK(ac)SAPATYK-Biotin; H3(13–32)K23ac, GKAPRKQLATK(ac)AARKSAPATYK-Biotin; H3(1–33)K4K23ac-Bio, ARTKQTARKSTGGKAPRKQLATK(ac)AARKSAPATGYK-Biotin; H3(1–33)K4-Bio, ARTKQTARKSTGGKAPRKQLATKAARKSAPATGYK-Biotin; H3(1–33)K4, ARTKQTARKSTGGKAPRKQLATKAARKSAPATG; H3K4K23ac, ARTKQTARKSTGGKAPRKQLATK(ac)AARKSAPATG; H3K4me3K23ac, ARTK(me3)QTARKSTGGKAPRKQLATK(ac)AARKSAPATG; H4(14–19)K16ac, GAK(ac)RHR; H4(1–20)K16ac, SGRGKGGKGLGKGGAK(ac)RHRK; H4(1–21), SGRGKGGKGLGKGGAKRHRKVGGK-Biotin; H4(1–27)K5ac, SGRGK(ac)GGKGLGKGGAKRHRKVLRDNIQ-PEG-Biotin; H4(1–27)K8ac, SGRGKGGK(ac)GLGKGGAKRHRKVLRDNIQ-PEG-Biotin; H4(1–27)K12ac, SGRGKGGKGLGK(ac)GGAKRHRKVLRDNIQ-PEG-Biotin; H4(1–27)K16ac, SGRGKGGKGLGKGGAK(ac)RHRKVLRDNIQ-PEG-Biotin; H4(1–27)K20ac, SGRGKGGKGLGKGGAKRHRK(ac)VLRDNIQ-PEG-Biotin.

44. Tsai, F. T. & Sigler, P. B. Structural basis of preinitiation complex assembly on human pol II promoters. *EMBO J.* **19**, 25–36 (2000).
45. Tsai, W. W., Nguyen, T. T., Shi, Y. & Barton, M. C. p53-targeted LSD1 functions in repression of chromatin structure and transcription in vivo. *Mol. Cell. Biol.* **28**, 5139–5146 (2008).
46. Bonéy-Montoya, J. *et al.* Long-range transcriptional control of progesterone receptor gene expression. *Mol. Endocrinol.* **24**, 346–358 (2010).
47. Kuhn, R. M. *et al.* The UCSC Genome Browser Database: update 2009. *Nucleic Acids Res.* **37** (Database issue), D755–D761 (2009).
48. Bentley, D. R. *et al.* Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**, 53–59 (2008).
49. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
50. Shi, X. *et al.* Proteome-wide analysis in *Saccharomyces cerevisiae* identifies several PHD fingers as novel direct and selective binding modules of histone H3 methylated at either lysine 4 or lysine 36. *J. Biol. Chem.* **282**, 2450–2455 (2007).
51. McCoy, A. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D* **63**, 32–41 (2007).
52. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
53. Brunger, A. T. Version 1.2 of the Crystallography and NMR system. *Nature Protocols* **2**, 2728–2733 (2007).
54. Jacobs, S. A., Fischle, W. & Khorasanizadeh, S. Assays for the determination of structure and dynamics of the interaction of the chromodomain with histone peptides. *Methods Enzymol.* **376**, 131–148 (2004).

# CRTC3 links catecholamine signalling to energy balance

Youngsup Song<sup>1</sup>, Judith Altarejos<sup>1</sup>, Mark O. Goodarzi<sup>2</sup>, Hiroshi Inoue<sup>1</sup>, Xiuqing Guo<sup>3</sup>, Rebecca Berdeaux<sup>1</sup>, Jeong-Ho Kim<sup>1</sup>, Jason Goode<sup>1</sup>, Motoyuki Igata<sup>1</sup>, Jose C. Paz<sup>1</sup>, Meghan F. Hogan<sup>1</sup>, Pankaj K. Singh<sup>1</sup>, Naomi Goebel<sup>1</sup>, Lili Vera<sup>1</sup>, Nina Miller<sup>1</sup>, Jinrui Cui<sup>3</sup>, Michelle R. Jones<sup>2</sup>, CHARGE Consortium, GIANT Consortium, Yii-Der I. Chen<sup>3</sup>, Kent D. Taylor<sup>3</sup>, Willa A. Hsueh<sup>4</sup>, Jerome I. Rotter<sup>3</sup> & Marc Montminy<sup>1\*</sup>

**The adipose-derived hormone leptin maintains energy balance in part through central nervous system-mediated increases in sympathetic outflow that enhance fat burning. Triggering of  $\beta$ -adrenergic receptors in adipocytes stimulates energy expenditure by cyclic AMP (cAMP)-dependent increases in lipolysis and fatty-acid oxidation. Although the mechanism is unclear, catecholamine signalling is thought to be disrupted in obesity, leading to the development of insulin resistance. Here we show that the cAMP response element binding (CREB) coactivator *Crtc3* promotes obesity by attenuating  $\beta$ -adrenergic receptor signalling in adipose tissue. *Crtc3* was activated in response to catecholamine signals, when it reduced adenylyl cyclase activity by upregulating the expression of *Rgs2*, a GTPase-activating protein that also inhibits adenylyl cyclase activity. As a common human *CRTC3* variant with increased transcriptional activity is associated with adiposity in two distinct Mexican-American cohorts, these results suggest that adipocyte *CRTC3* may play a role in the development of obesity in humans.**

Obesity is a major risk factor in the development of insulin resistance, which is characterized by decreased glucose uptake into muscle and increased glucose production by the liver. Obesity affects one-third of adults in the USA<sup>1</sup>; the prevalence of obesity appears even higher in certain ethnic groups, although relevant predisposing factors have not been fully identified. Obesity is a particular problem among Mexican-Americans, with an overall prevalence of 40% (36% in men, 45% in women)<sup>1</sup>, contributing to elevated rates of diabetes<sup>2</sup>. Environment, lifestyle and genetic susceptibility probably contribute to the increased risk of obesity and diabetes in this population.

Under lean conditions, the adipose-derived hormone leptin is thought to promote energy expenditure through increases in sympathetic nerve activity that enhance catecholamine signalling in white adipose tissue (WAT) and brown adipose tissue (BAT)<sup>3,4</sup>. Triggering of  $\beta$ -adrenergic receptors appears important for subsequent increases in lipolysis and fatty-acid oxidation<sup>5</sup>; mice with knockouts of all three  $\beta$  receptors have reduced energy expenditure, and they are more susceptible to effects of high-fat diet (HFD) feeding on weight gain. Conversely, transgenic over-expression of  $\beta$ -adrenergic receptor 1 in adipose tissue appears sufficient to confer resistance to obesity<sup>6</sup>.

Triggering of  $\beta$ -adrenergic receptors stimulates cAMP-mediated increases in cellular gene expression with burst-attenuation kinetics<sup>7,8</sup>; rates of transcription peak within 1 h of stimulation and decrease thereafter even under continuous stimulation. Although the underlying mechanisms remain unclear, the attenuation of cellular genes is thought to be coordinated by negative feedback effectors, which are themselves targets for upregulation by cAMP<sup>9</sup>.

cAMP stimulates the expression of cellular genes through the protein kinase A (PKA)-mediated phosphorylation of CREB family members (CREB1, ATF1, CREM), a modification that promotes recruitment of the histone acetyl transferase paralogues P300 and CBP<sup>10–12</sup>.

In parallel, cAMP also increases gene expression by stimulating the CREB regulated transcriptional coactivators (CRTC3)<sup>13,14</sup>. Under basal

conditions, CRTCs are sequestered in the cytoplasm through phosphorylation-dependent interactions with 14-3-3 proteins. CRTCs are phosphorylated by salt-inducible kinases and other members of the stress- and energy-sensing AMPK family of Ser/Thr kinases. Increases in intracellular cAMP signalling promote the PKA-mediated phosphorylation and inhibition of salt-inducible kinase activity, leading to the subsequent dephosphorylation and nuclear entry of CRTCs, which bind to CREB over relevant promoters. After prolonged stimulation with cAMP agonist, CRTC activity is terminated through ubiquitin-mediated degradation<sup>15</sup>.

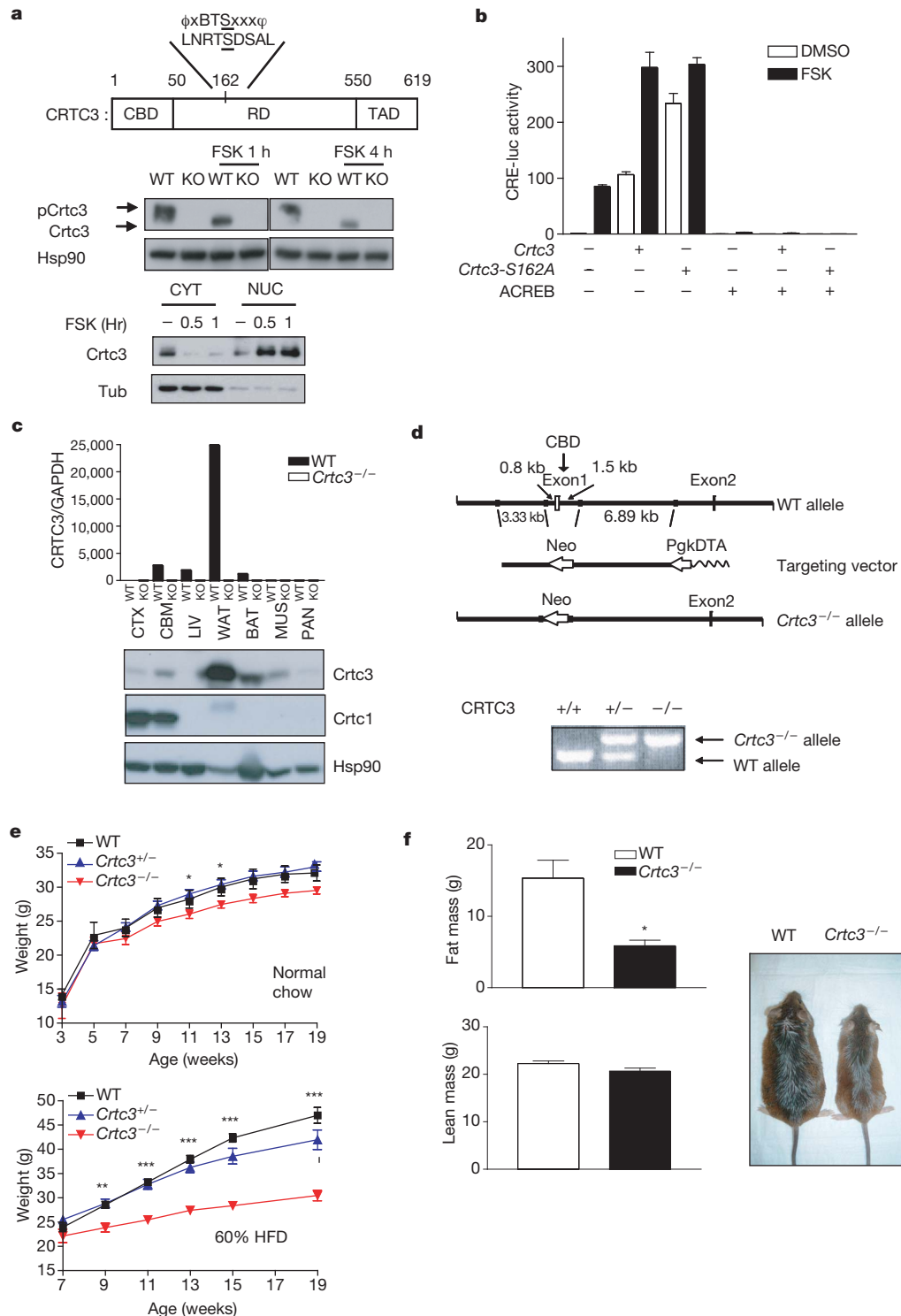
The CRTC family consists of three members (*Crtc1*, *Crtc2* and *Crtc3*), which are distinguished in part by their expression profiles. *Crtc1* is produced primarily in brain, where it mediates leptin effects on satiety<sup>16</sup>; mice with a knockout of the *Crtc1* gene develop obesity due in part to reductions in energy expenditure. By contrast, *Crtc2* is expressed at high levels in liver where it promotes fasting gluconeogenesis<sup>17,18</sup>; mice with a knockout of *Crtc2* appear more insulin sensitive under HFD conditions, owing to reductions in hepatic glucose output<sup>19</sup>.

## Role of CRTC3 as a CREB coactivator

Similar to other CRTC family members, *Crtc3* contains CREB binding (CBD; amino acids 1–50), regulatory (RD; amino acids 51–549) and trans-activation domains (TAD; amino acids 550–619), which are also present in *Crtc1* and *Crtc2* (Fig. 1a). In the basal state, *Crtc3* is phosphorylated at Ser 162 by salt-inducible kinases and other members of the stress- and energy-sensing AMPK family of Ser/Thr kinases<sup>13,20,21</sup>. Short-term (0.5–1 h) exposure to cAMP agonist promotes the dephosphorylation and nuclear entry of *Crtc3* (Fig. 1a); similar to *Crtc2*<sup>15</sup>, prolonged cAMP stimulation triggers *Crtc3* degradation.

*Crtc3* over-expression augments the activity of a cAMP responsive (CRE-luc) reporter in cells exposed to forskolin (FSK; Fig. 1b); and mutation of the regulatory Ser162 phosphorylation site to alanine further enhances *Crtc3* activity under basal conditions. In keeping with the

<sup>1</sup>The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, USA. <sup>2</sup>Division of Endocrinology, Diabetes and Metabolism, Department of Medicine, Cedars-Sinai Medical Center, Los Angeles, California 90048, USA. <sup>3</sup>Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, California 90048, USA. <sup>4</sup>Diabetes Research Center, Division of Diabetes, Obesity and Lipids, Methodist Hospital Research Institute, Houston, Texas 77030, USA.



**Figure 1 | *Crtc3*<sup>-/-</sup> mice are resistant to obesity.** **a**, Top, CREB-binding (CBD), regulatory (RD) and transactivation (TAD) domains and conserved AMPK/salt-inducible kinase phosphorylation site (Ser 162). Consensus phosphorylation site for AMPK family members ( $\phi$ xBTSTxxx $\phi$ ) shown; relative position of hydrophobic ( $\phi$ ), basic (B), Thr (T), and phosphorylated Ser (S) residues indicated. x represents any amino acid. Middle, immunoblot of Crtc3 in wild-type (WT) and *Crtc3*<sup>-/-</sup> (knockout, KO) MEFs exposed to FSK. Bottom, effect of FSK on nuclear and cytoplasmic Crtc3 levels. **b**, Effect of wild-type or S162A Crtc3 on CRE-luciferase activity. **c**, Quantitative PCR (top) and

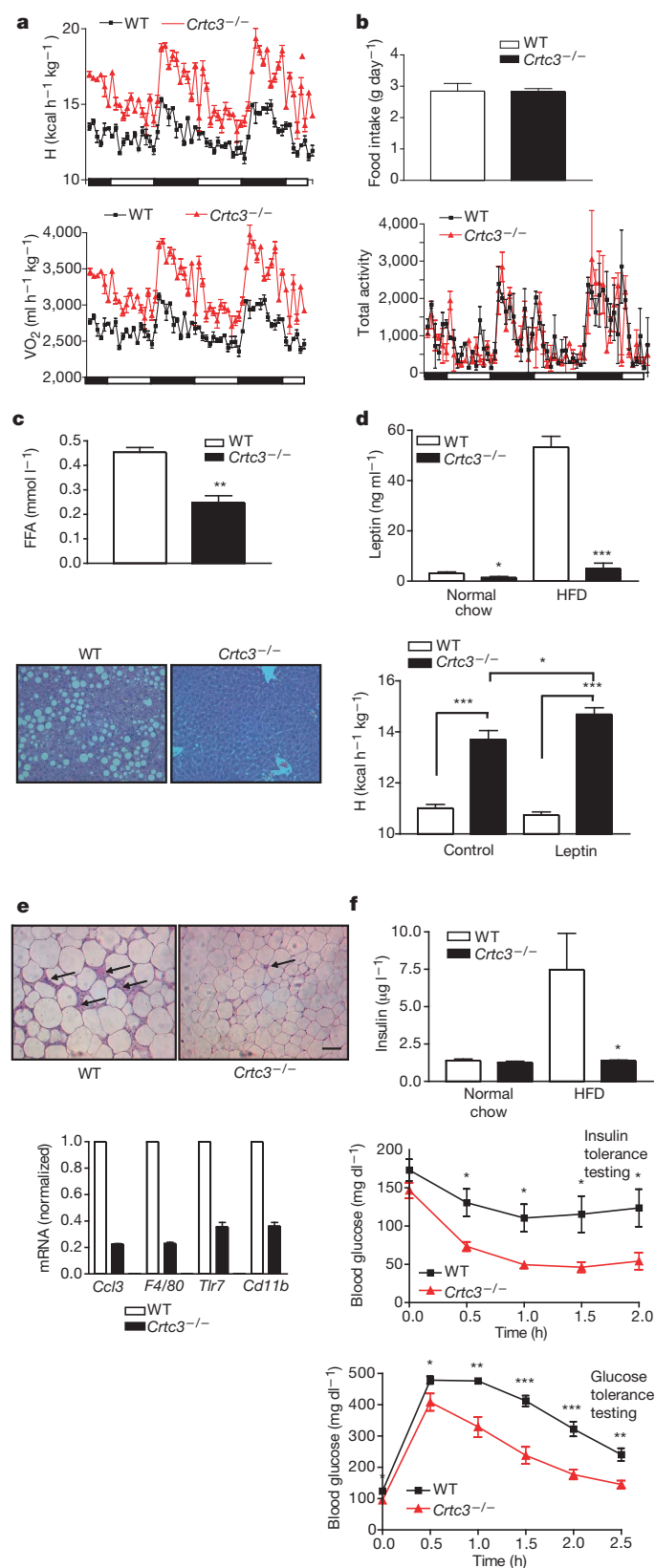
immunoblot (bottom) analysis of Crtc3 tissue expression. BAT, brown adipose; CBM, cerebellum; CTX, cortex; LIV, liver; MUS, skeletal muscle; PAN, pancreas; WAT, white adipose. **d**, Top, *Crtc3* targeting vector with Neo selection marker replacing Exon 1, which encodes the CBD. Bottom, PCR analysis of wild-type and mutant *Crtc3* alleles in mice. **e**, Weight gain in wild-type and *Crtc3* mutant mice maintained on normal chow ( $n=8$  per group) or HFD ( $n=5$ ) (\* $P<0.05$ ; \*\* $P<0.01$ ; \*\*\* $P<0.001$ ). **f**, Fat mass (left) and photograph (right) of HFD-fed wild-type and *Crtc3*<sup>-/-</sup> mice ( $n=4$  per group) (\* $P<0.05$ ). Error bars, s.e.m.

proposed role of CREB in recruiting CRTC3 to relevant promoters, expression of a dominant negative CREB inhibitor, called ACREB<sup>22</sup>, blocks Crtc3 effects on reporter activity in cells exposed to FSK. By

contrast with Crtc1, which is expressed primarily in brain, Crtc3 protein and messenger RNA (mRNA) amounts are particularly abundant in WAT and to a lesser extent in BAT (Supplementary Fig. 1 and Fig. 1c).



Based on the importance of the CBD for *Crtc3*-mediated induction of cAMP-responsive genes<sup>23,24</sup>, we generated *Crtc3*<sup>-/-</sup> mice with a deletion of exon 1, which encodes the CBD (Fig. 1d). *Crtc3*<sup>-/-</sup> mice are born at the expected Mendelian frequency; they appear comparable to wild-type littermates at birth, despite the absence of detectable *Crtc3* mRNA and protein amounts in all tissues (Fig. 1c).



## Role of CRT3 in energy balance

When maintained on a normal chow diet, *Crtc3*<sup>-/-</sup> mice appear more insulin sensitive than controls by insulin tolerance testing (Supplementary Fig. 1, right). *Crtc3*<sup>-/-</sup> animals also have 50% lower adipose tissue mass, despite comparable food intake and physical activity to control mice (Supplementary Fig. 2).

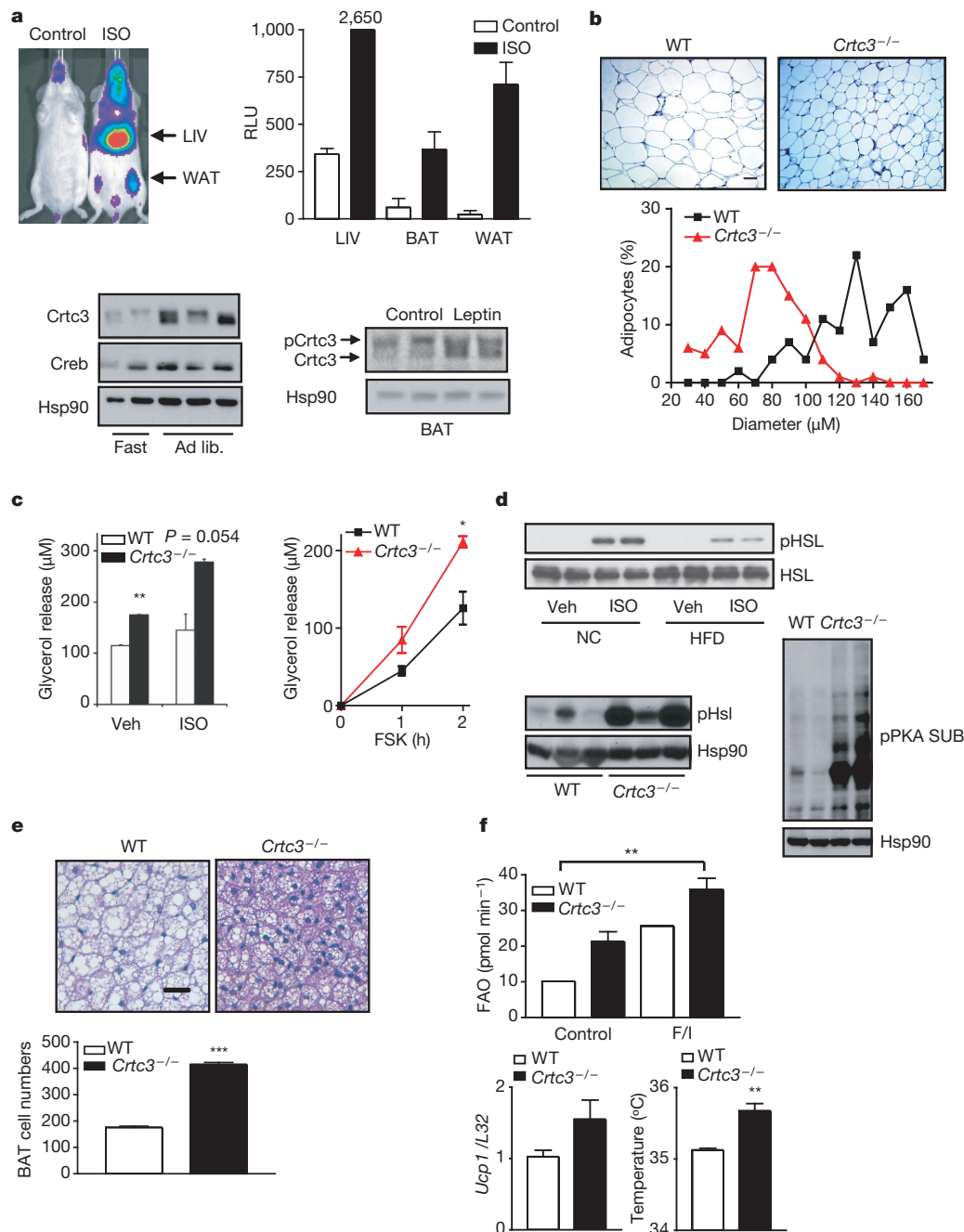
When transferred to an HFD (60% of calories from fat), *Crtc3*<sup>-/-</sup> mice gained 35% less weight relative to controls reflecting primarily differences in fat accumulation (Fig. 1e, f). The effect of *Crtc3* on adiposity appeared to be dependent on gene dosage as *Crtc3*<sup>+/-</sup> mice show intermediate weight gains relative to wild-type and *Crtc3*<sup>-/-</sup> mice. Although physical activity and food intake were nearly identical, energy expenditure and oxygen consumption were substantially elevated in HFD-fed *Crtc3*<sup>-/-</sup> mice relative to wild-type littermates (Fig. 2a, b). Pointing to parallel increases in glucose and lipid oxidation, respiratory quotients were comparable in wild-type and *Crtc3*<sup>-/-</sup> mice (Supplementary Fig. 3).

Circulating concentrations of free fatty acids were decreased in *Crtc3*<sup>-/-</sup> mice, and they were protected from the effects of HFD feeding on hepatic steatosis (Fig. 2c). Consistent with their reduced fat mass, *Crtc3*<sup>-/-</sup> mice had decreased circulating leptin concentrations compared with wild-type littermates, although the reduction in leptin levels (tenfold) appeared disproportionately low relative to the difference in fat mass (threefold) (Fig. 2d and Supplementary Fig. 4). Indeed, intraperitoneal administration of leptin stimulated energy expenditure to a greater extent in *Crtc3* mutant than wild-type mice. Taken together, these results indicate that disruption of *Crtc3* activity leads to increases in energy expenditure, which maintain leptin sensitivity and protect against ectopic lipid accumulation.

Under obese conditions, increases in inflammatory infiltrates in adipose tissue contribute to the development of systemic insulin resistance<sup>25</sup>. Although they were readily observed in wild-type mice, adipose-tissue macrophages were less abundant in *Crtc3*<sup>-/-</sup> tissue (Fig. 2e and Supplementary Fig. 5). Arguing against an effect of the *Crtc3* knockout on macrophage function per se, tumour necrosis factor- $\alpha$  release from peritoneal macrophages in response to lipopolysaccharide appeared comparable between *Crtc3* mutant and control cells (Supplementary Fig. 5). In line with these differences, circulating insulin concentrations were lower in HFD-fed *Crtc3*<sup>-/-</sup> than wild-type mice, and whole-body insulin sensitivity was correspondingly improved by insulin and glucose tolerance testing (Fig. 2f). As a result, glucose uptake into muscle was increased in *Crtc3*<sup>-/-</sup> mice compared with control littermates (Supplementary Fig. 6).

We considered that *Crtc3* activity in adipose tissue may be modulated by hormonal signals. In line with its effects on *Crtc3* dephosphorylation in cell cultures (Supplementary Fig. 7), intraperitoneal administration of  $\beta$ -adrenergic agonist isoproterenol (ISO) increased the activity of a CRE-luc reporter transgene in WAT and BAT by live imaging analysis (Fig. 3a). Leptin administration (intraperitoneal) also promoted *Crtc3* dephosphorylation. *Crtc3* protein amounts in WAT are elevated under *ad libitum* conditions; they decreased after fasting for 6, when *Crtc3* appeared to undergo degradation (Fig. 3a). Consistent with an increase in protein stability under obese conditions, Creb and *Crtc3* protein amounts were upregulated in WAT from HFD-fed mice compared with those fed on normal chow (Supplementary Fig. 8).

**Figure 2 | Increased energy expenditure in *Crtc3*<sup>-/-</sup> mice.** **a, b**, Energy expenditure and oxygen consumption (**a**) as well as food intake and physical activity (**b**) in HFD-fed mice ( $n=4$  per group). **c**, Free-fatty-acid levels (top) and haematoxylin and eosin sections of livers (bottom) in HFD-fed mice ( $n=3$  per group) ( $**P<0.01$ ). **d**, Leptin levels (top) ( $n=5$  per group) and effect of intraperitoneal leptin administration on energy expenditure (bottom) ( $n=4$  per group) ( $*P<0.05$ ;  $***P<0.001$ ). **e**, Macrophage infiltration (top) and gene expression (bottom) in WAT from HFD-fed mice. Scale bar, 50  $\mu\text{m}$ . **f**, Insulin levels (top), insulin tolerance testing (middle) and glucose tolerance testing (bottom) of HFD-fed mice ( $n=5$  per group) ( $*P<0.05$ ;  $**P<0.01$ ;  $***P<0.001$ ). Error bars, s.e.m.



**Figure 3 | Increased catecholamine signalling in *Crtc3*<sup>-/-</sup> adipose tissue.** **a**, Top, effect of ISO on CRE-luc reporter activity in different tissues. Bottom, immunoblots of *Crtc3* in WAT (left) or BAT (right). Effect of fasting (Fast), feeding *ad libitum* (Ad lib.) and leptin administration indicated. LIV, liver. **b**, Haematoxylin and eosin sections (top) and adipocyte size distribution (bottom) in WAT from wild-type and *Crtc3*<sup>-/-</sup> mice. Scale bar, 50  $\mu$ m. **c**, Lipolysis rates in adipocytes exposed to ISO (left) or FSK (right) ( $n=3$ ) (\* $P<0.05$ ; \*\* $P<0.01$ ). Veh, vehicle. **d**, Phospho- (Ser 660) HSL levels in WAT

from mice fed on normal chow or HFD after ISO injection (top) and from HFD-fed wild-type or *Crtc3*<sup>-/-</sup> mice (bottom left). Bottom right, immunoblot of PKA activity in WAT from HFD-fed mice. **e**, Haematoxylin and eosin sections (top) and brown adipocyte numbers (bottom) in wild-type and *Crtc3*<sup>-/-</sup> BAT. Scale bar, 50  $\mu$ m (\*\* $P<0.001$ ). **f**, Top, fatty-acid oxidation (FAO) and *Ucp1* mRNA levels in brown adipocytes. Core body temperatures indicated ( $n=4$  per group) (\*\* $P<0.01$ ). Error bars, s.e.m. F/I, exposure to forskolin plus isoproterenol.

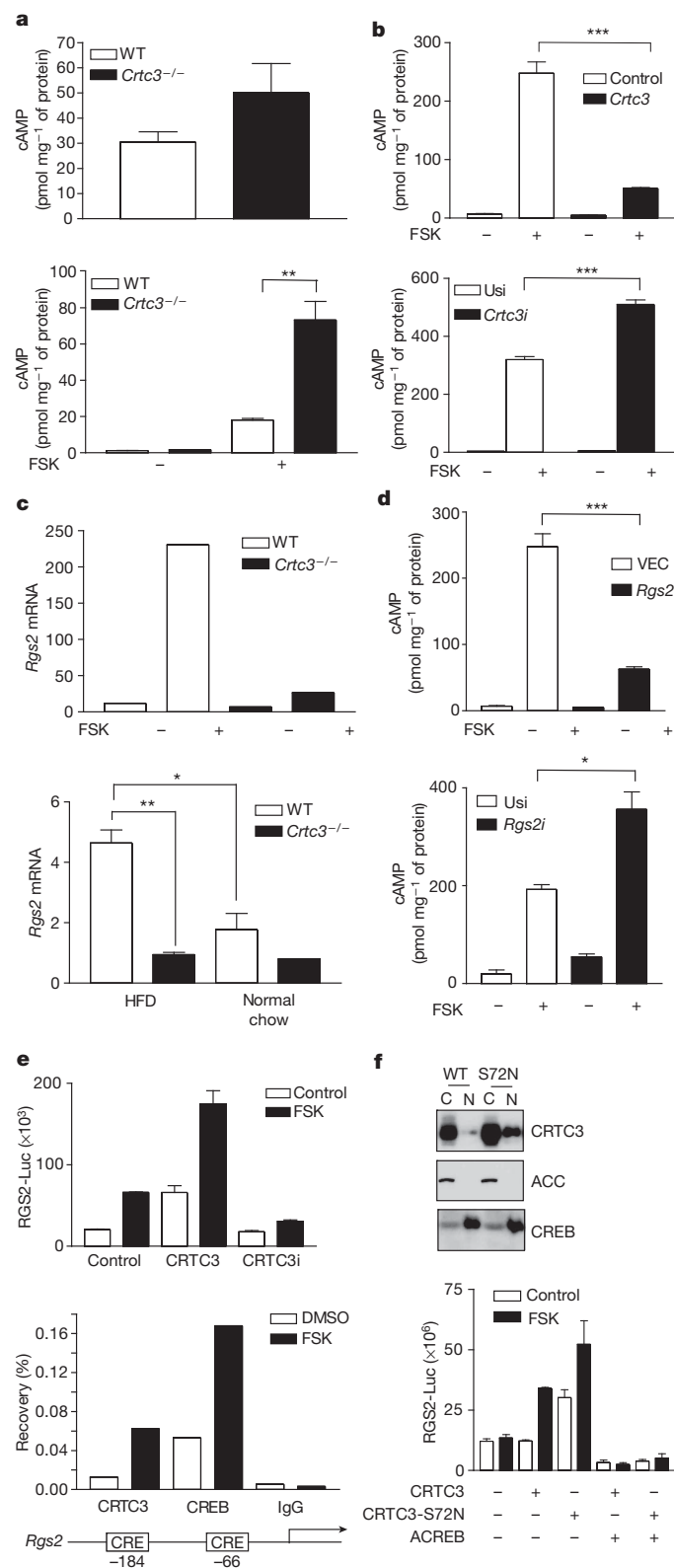
### Catecholamine signalling in adipose tissue

Under HFD feeding conditions, increases in catecholamine signalling maintain energy balance by mobilizing triglyceride stores in WAT<sup>26</sup>. Although the total number of adipocytes in WAT fat pads was nearly identical in both groups, adipocytes from *Crtc3*<sup>-/-</sup> mice were substantially smaller than from wild-type mice (Fig. 3b). Arguing against a disruption in triglyceride synthesis, mRNA amounts for lipogenic genes (*Acc*, *Lpl*, *Scd*) appeared comparable between *Crtc3* mutant and wild-type adipocytes (Supplementary Fig. 9). Rather, basal and ISO-induced lipolysis rates were increased in *Crtc3*<sup>-/-</sup> compared with

control adipocytes (Fig. 3c). Exposure to FSK also increased lipolysis to a greater extent in *Crtc3*<sup>-/-</sup> adipocytes (Fig. 3c), pointing to the potential upregulation of the cAMP signalling pathway in these cells.

Triggering of  $\beta$ -adrenergic receptors has been found to promote lipolysis through the cAMP-dependent PKA-mediated phosphorylation of hormone sensitive lipase (HSL)<sup>27</sup>. In keeping with the proposed downregulation of  $\beta$ -adrenergic receptor signalling in obesity, administration of ISO had only modest effects on HSL phosphorylation in HFD-fed relative to animals fed on normal chow (Fig. 3d). Indeed, amounts of phospho- (Ser 660) HSL were substantially elevated

in *Crtc3*<sup>-/-</sup> WAT compared with wild type, even though circulating concentrations of noradrenaline and adrenaline were similar between the two groups (Fig 3d and Supplementary Fig. 10). PKA activity in WAT was also increased in *Crtc3*<sup>-/-</sup> mice by immunoblot assay using a phospho-specific PKA substrate antiserum (Fig. 3d). Consistent with the predominant expression of *Crtc3* in adipose tissue, PKA activity in other tissues appeared similar between wild-type and *Crtc3*<sup>-/-</sup> mice (Supplementary Fig. 11).



Having seen that lipolysis rates are increased in WAT, and realizing that circulating free-fatty-acid concentrations are reduced in *Crtc3* mutant mice, we considered that fatty-acid oxidation should also be upregulated in this setting. Under HFD conditions, leptin has been proposed to trigger catecholamine-mediated increases in fat burning in BAT, a process known as diet-induced thermogenesis<sup>28,29</sup>. In keeping with the ability for catecholamines to stimulate BAT expansion, brown adipocyte numbers were increased twofold in intra-scapular fat pads from *Crtc3*<sup>-/-</sup> mice compared with controls (Fig. 3e). Suggesting a parallel increase in fat burning, *Crtc3*<sup>-/-</sup> brown adipocytes also had smaller intracellular lipid vacuoles than wild-type cells. Moreover, fatty-acid oxidation rates were increased in primary brown adipocytes from *Crtc3*<sup>-/-</sup> mice relative to controls, and uncoupling protein 1 (*Ucp1*) mRNA amounts were also higher (Fig. 3f). Correspondingly, core body temperatures were elevated in *Crtc3*<sup>-/-</sup> mice compared with control animals. Taken together, these results indicate that loss of *Crtc3* increases fat burning in part through increases in brown adipocyte numbers in BAT.

We reasoned that the loss of *Crtc3* expression could increase cellular PKA activity by altering the subunit composition of the PKA holoenzyme. Over-expression of the forkhead protein *Foxc2* in WAT, for example, has been found to promote energy expenditure by upregulating mRNA amounts for regulatory subunit I, which has a higher affinity for cAMP than RII<sup>30</sup>. Arguing against this possibility, however, mRNA amounts for regulatory subunits I and II in WAT were comparable between wild type and *Crtc3* mutants (Supplementary Fig. 12, left).

### cAMP accumulation in CRTC3 mutant cells

Alternatively, disruption of the *Crtc3* gene may enhance PKA activity by increasing cellular cAMP accumulation in response to hormonal signals. Supporting this idea, cAMP concentrations were elevated in WAT from *Crtc3* mutant mice relative to controls (Fig. 4a). Exposure to FSK also triggered cAMP accumulation to a greater extent in *Crtc3*<sup>-/-</sup> mouse embryonic fibroblasts (MEFs) and in *Crtc3*<sup>-/-</sup> BAT stromal-vascular cells compared with wild-type cells (Fig. 4a and Supplementary Fig. 12, right). Moreover, *Crtc3* over-expression in wild-type cells reduced cAMP production in response to FSK, whereas acute RNA interference (RNAi)-mediated depletion of *Crtc3* increased it (Fig. 4b and Supplementary Fig. 13).

In principle, the enhanced accumulation of cAMP in *Crtc3*-deficient cells could reflect a decrease in cellular phosphodiesterase activity. In that event, treatment with non-selective phosphodiesterase inhibitor should lead to comparable increases in cAMP concentrations between wild-type and mutant cells exposed to  $\beta$ -adrenergic agonist. However, intracellular cAMP concentrations remained higher in *Crtc3*<sup>-/-</sup> than wild-type cells after co-stimulation with ISO plus isobutyl-methyl xanthine (Supplementary Fig. 14). Based on these results, we reasoned that *Crtc3* probably inhibits cAMP signalling by modulating cellular adenylyl cyclase activity.

In gene profiling studies to identify cellular genes that mediate inhibitory effects of *Crtc3* on cAMP signalling, we identified the Regulator of G protein signalling 2 (*Rgs2*) as the most highly upregulated gene of 15 that are induced threefold or better in adipocytes exposed to FSK<sup>31</sup>. We confirmed these effects in cultured primary

**Figure 4 | *Crtc3* attenuates adipose tissue cAMP signalling.** **a**, cAMP content in WAT (top) and MEFs (bottom) from wild-type and *Crtc3*<sup>-/-</sup> mice (\*\*P < 0.01). **b**, cAMP accumulation in MEFs over-expressing (top) or depleted of (bottom) *Crtc3*. (\*\*\*P < 0.001). **c**, *Rgs2* mRNA levels in cultured adipocytes exposed to FSK (top) and in WAT from mice fed on normal chow or HFD (bottom) (\*P < 0.05; \*\*P < 0.01). **d**, Effect of *Rgs2* over-expression (top) or *Rgs2* RNAi (bottom) on cAMP accumulation in MEFs (\*P < 0.05; \*\*\*P < 0.001). **e**, Top, effect of *Crtc3* over-expression or *Crtc3* RNAi on RGS2-luc reporter activity. Bottom, chromatin immunoprecipitation assay of *Crtc3* occupancy over the *Rgs2* promoter. **f**, Top, relative nuclear/cytoplasmic fractionation (top) and activities (bottom) of over-expressed wild-type and S72N CRTC3 in HEK293T cells. Error bars, s.e.m.



adipocytes, where exposure to FSK increased the expression of *Rgs2* and other targets in wild-type cells, but to a much lesser extent in *Crtc3*<sup>-/-</sup> cells (Fig. 4c, top, and Supplementary Fig. 15). By contrast with its reduction in adipose tissue from *Crtc3*<sup>-/-</sup> mice, *Creb* target gene expression in skeletal muscle appeared comparable between wild-type and *Crtc3* mutant animals, probably reflecting regulatory contributions from *Crtc2*, which has been shown to promote the expression of PGC-1 $\alpha$  and mitochondrial genes in muscle cells<sup>32</sup>.

First identified as a GTPase activating protein that blocks Gq signalling, *Rgs2* has also been shown to inhibit the cAMP pathway by binding directly to a subset of adenylyl cyclases (types III, V and VI)<sup>33–35</sup>, isoforms that are enriched in BAT and WAT<sup>36,37</sup>. Moreover, mutations that enhance *RGS2* expression have been associated with increased risk of metabolic syndrome in humans<sup>38</sup>. In keeping with the upregulation of *Crtc3* in obesity, HFD feeding stimulated *Rgs2* mRNA amounts in wild-type WAT, but *Rgs2* expression remained low in WAT from HFD-fed *Crtc3*<sup>-/-</sup> mice (Fig. 4c, bottom). Consistent with its proposed role as an adenylyl cyclase inhibitor, *Rgs2* over-expression reduced cAMP production in cells exposed to FSK, whereas RNAi-mediated knockdown of *Rgs2* increased it (Fig. 4d and Supplementary Figs 16 and 17).

We examined whether the *Rgs2* gene is a direct target of *Creb* and *Crtc3*. In line with the presence of conserved CREB binding sites at -184 and -66 on the *Rgs2* promoter, exposure to FSK upregulated *RGS2*-luciferase reporter activity in transient transfection assays (Fig. 4e). *Crtc3* over-expression further enhanced *RGS2* promoter activity, whereas RNAi-mediated depletion of *Crtc3* reduced it. Consistent with a direct effect of these activators, exposure to FSK increased *Creb* and *Crtc3* occupancy over the *RGS2* promoter in wild-type cells (Fig. 4e).

## Role of *CRTC3* in human obesity

Having seen the effects of *Crtc3* on energy expenditure, we wondered whether this coactivator also contributes to obesity in humans. Within the human database of single nucleotide polymorphisms (dbSNP), we noticed a common *CRTC3* variant allele, which encodes a missense variant (S72N) near a predicted nuclear export sequence<sup>39</sup>. Supporting this idea, nuclear amounts of 72N *CRTC3* were elevated relative to 72S *CRTC3* under basal conditions (Fig. 4f). Correspondingly, 72N variant *CRTC3* was more potent than 72S *CRTC3* in stimulating *RGS2* promoter activity, particularly under basal conditions (Fig. 4f).

We examined the potential association between the S72N variant *CRTC3* and adiposity in a Mexican-American cohort of 779 individuals (Table 1). The allele frequency of the 72N variant in this population was 34%. In keeping with its increased activity relative to wild-type *CRTC3*, the 72N allele was also associated with several anthropometric indices of adiposity including weight, body mass index (BMI) and hip circumference. Similar to the gene dosage effect of *CRTC3* on weight gain in mice, Mexican-Americans with two 72N alleles had increased adiposity compared with those with only one variant allele; and 72S/72N heterozygous individuals had intermediate adiposity relative to individuals homozygous for the wild-type and variant *CRTC3* alleles.

We then sought to confirm the association of 72N with increases in adiposity indices, by assessing the association of a perfect proxy SNP

**Table 1 | Association of S72N with anthropometric indices in the MACAD cohort**

	S/S (n=346)	S/N (n=338)	N/N (n=95)	P value
Weight (kg)	73.3 (19.6)	75.1 (20.6)	76.0 (22.2)	0.033*
BMI (kg/m <sup>2</sup> )	28.1 (6.1)	28.3 (6.1)	29.1 (5.6)	0.038*
Hip circumference (cm)	103.0 (12.3)	104.5 (13.0)	104.5 (12.8)	0.033*
Waist (cm)	90.5 (15)	93.0 (16)	93.3 (17.3)	0.15
BSA (m <sup>2</sup> )	1.77 (0.28)	1.81 (0.30)	1.84 (0.30)	0.075
Waist/hip ratio	0.88 (0.10)	0.89 (0.11)	0.88 (0.12)	0.43

Values are median (interquartile range). S/S, individuals homozygous for wild-type *CRTC3* (Ser 72); S/N, individuals heterozygous for variant *CRTC3* (Asn 72); N/N, individuals homozygous for variant (Asn 72) *CRTC3*.

\* Significant P values.

rs3862434 ( $r^2=1$  with S72N) with adiposity measures in 987 Mexican-Americans from the Multi-Ethnic Study of Atherosclerosis (MESA). The minor allele of rs3862434 (G, frequency 34%), which corresponds to 72N, was associated with increased body surface area (BSA) and a trend to increased weight (Supplementary Table 1). In an analysis combining the two cohorts, 72N exhibited associations with weight, hip circumference, BMI and BSA (Supplementary Table 2). Because these traits are interrelated, we did not use multiple testing correction, which would be overly conservative because our other data suggested 72N altered the biological function of *CRTC3*. Taken together, these results indicate that *CRTC3* is associated with increased obesity risk in Mexican-Americans.

This association did not extend to 2,527 non-Hispanic whites from MESA, in whom we observed no association of rs3862434 with any of the six available obesity traits (Supplementary Table 3). Of note, the G allele of rs3862434 (proxy for the adiposity-associated 72N allele) is the minor allele in Mexican-Americans and the major allele in non-Hispanic whites. We then conducted a z-score meta-analysis in over 63,000 subjects from the Cohorts for Heart and Aging Research in Genome Epidemiology (CHARGE) and Genetic Investigation of Anthropometric Traits (GIANT) consortia, examining the association of S72N with BMI. In this meta-analysis, the 72N allele was associated with increased BMI; however, this did not reach statistical significance (z-score 0.74,  $P=0.45$ ). This suggests that the effect of 72N to promote obesity, although substantial in Mexican-Americans, is very weak or non-existent in non-Hispanic whites. As the 72N allele is more frequent in non-Hispanic whites, and yet has a minimal effect on BMI in this population, our results are consistent with the observation that obesity is less frequent in non-Hispanic whites than Mexican-Americans<sup>1</sup>. The weaker effect of 72N in non-Hispanic whites may be due to environmental/lifestyle factors and/or differences in genetic background.

## Discussion

HFD feeding has been shown to promote obesity and insulin resistance through increases in energy intake that lead to the ectopic deposition of lipid in liver. Our results suggest that *Crtc3* contributes to these changes in part by attenuating catecholamine signalling in adipose tissue (Supplementary Fig. 18).

Although the proposed role of *Crtc3* as a negative feedback regulator of adipocyte cAMP signalling was unexpected, we note that intracellular signalling pathways often self-attenuate as part of a homeostatic mechanism to limit cellular responses to hormonal stimuli<sup>7</sup>. Thus the chronic upregulation of sympathetic nerve activity under HFD conditions<sup>26</sup> may attenuate the intracellular cAMP pathway through the *Crtc3*-mediated induction of *Rgs2*. By limiting catecholamine-dependent increases in lipolysis and fatty-acid oxidation, *Crtc3* may also function as a so-called 'thrifty' gene that enhances survival under starvation conditions. Future studies may reveal the extent to which *CRTC3* also contributes to the development of insulin resistance and type II diabetes.

## METHODS SUMMARY

*Crtc3*<sup>-/-</sup> mice were generated by targeted disruption of exon 1 in the *Crtc3*<sup>-/-</sup> gene, which encodes the CREB binding domain. Transgenic CRE-luciferase reporter mice were generated and analysed by *in vivo* imaging with an IVIS-100 instrument (Caliper Life Sciences). For indirect calorimetry studies, mice were housed individually; oxygen consumption and energy expenditure were measured using a LabMaster system (TSE Systems). Circulating leptin, insulin and free-fatty-acid concentrations were measured by enzyme-linked immunosorbent assay (ELISA). *In vitro* lipolysis and fatty oxidation rates were evaluated on cultures of primary white and brown adipocytes as described<sup>31,40</sup>; glucose uptake was measured in soleus muscle using [U-<sup>14</sup>C]2-deoxyglucose. PKA activity was examined in different tissues by immunoblot assay with phospho-PKA substrate antiserum (Cell Signaling). Intracellular cAMP accumulation in wild-type and *CRTC3* mutant cells was measured by ELISA. *Crtc3* and *Creb* occupancies over cellular genes were evaluated by chromatin immunoprecipitation assay<sup>13</sup>. For studies of human genetic association, that of *CRTC3* variant (S72N) with

adiposity parameters was assessed in participants of the Cedars-Sinai/University of California, Los Angeles, Mexican-American Coronary Artery Disease (MACAD) study, comprising 779 subjects<sup>41,42</sup>. Confirmatory studies were performed on Mexican-American subjects from the Multi-Ethnic Study of Atherosclerosis (MESA)<sup>43</sup>, containing 987 Mexican-American individuals. Genotyping of SNP rs8033595 (S72N) in MACAD was performed using TaqManMGB technology<sup>42,44</sup>. This SNP was also explored for association with body mass index in the CHARGE and GIANT consortia.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 29 January; accepted 11 October 2010.**

- Flegal, K. M., Carroll, M. D., Ogden, C. L. & Curtin, L. R. Prevalence and trends in obesity among US adults, 1999–2008. *J. Am. Med. Assoc.* **303**, 235–241 (2010).
- Cowie, C. C. *et al.* Full accounting of diabetes and pre-diabetes in the U.S. population in 1988–1994 and 2005–2006. *Diabetes Care* **32**, 287–294 (2009).
- Bartness, T. J. & Song, C. K. Thematic review series: adipocyte biology. Sympathetic and sensory innervation of white adipose tissue. *J. Lipid Res.* **48**, 1655–1672 (2007).
- Lafontan, M. & Langin, D. Lipolysis and lipid mobilization in human adipose tissue. *Prog. Lipid Res.* **48**, 275–297 (2009).
- Bachman, E. S. *et al.*  $\beta$ AR signaling required for diet-induced thermogenesis and obesity resistance. *Science* **297**, 843–845 (2002).
- Soloveva, V., Graves, R. A., Rasenick, M. M., Spiegelman, B. M. & Ross, S. R. Transgenic mice overexpressing the  $\beta$ 1-adrenergic receptor in adipose tissue are resistant to obesity. *Mol. Endocrinol.* **11**, 27–38 (1997).
- Hagiwara, M. *et al.* Transcriptional attenuation following cAMP induction requires PP-1-mediated dephosphorylation of CREB. *Cell* **70**, 105–113 (1992).
- Michael, L. F., Asahara, H., Shulman, A., Kraus, W. & Montminy, M. The phosphorylation status of a cyclic AMP-responsive activator is modulated via a chromatin-dependent mechanism. *Mol. Cell. Biol.* **20**, 1596–1603 (2000).
- De Cesare, D. & Sassone-Corsi, P. Transcriptional regulation by cyclic AMP-responsive factors. *Prog. Nucleic Acid Res. Mol. Biol.* **64**, 343–369 (2000).
- Gonzalez, G. A. & Montminy, M. R. Cyclic AMP stimulates somatostatin gene transcription by phosphorylation of CREB at Serine 133. *Cell* **59**, 675–680 (1989).
- Chrivia, J. C. *et al.* Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature* **365**, 855–859 (1993).
- Radhakrishnan, I. *et al.* Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator–coactivator interactions. *Cell* **91**, 741–752 (1997).
- Screaton, R. A. *et al.* The CREB coactivator TORC2 functions as a calcium- and cAMP-sensitive coincidence detector. *Cell* **119**, 61–74 (2004).
- Bittinger, M. A. *et al.* Activation of cAMP response element-mediated gene expression by regulated nuclear transport of TORC proteins. *Curr. Biol.* **14**, 2156–2161 (2004).
- Liu, Y. *et al.* A fasting inducible switch modulates gluconeogenesis via activator/coactivator exchange. *Nature* **456**, 269–273 (2008).
- Altarejos, J. Y. *et al.* The Creb1 coactivator Crtc1 is required for energy balance and fertility. *Nature Med.* **14**, 1112–1117 (2008).
- Koo, S. H. *et al.* The CREB coactivator TORC2 is a key regulator of fasting glucose metabolism. *Nature* **437**, 1109–1111 (2005).
- Saber, M. *et al.* Novel liver-specific TORC2 siRNA corrects hyperglycemia in rodent models of type 2 diabetes. *Am. J. Physiol. Endocrinol. Metab.* **297**, E1137–E1146 (2009).
- Wang, Y. *et al.* Targeted disruption of the CREB coactivator Crtc2 increases insulin sensitivity. *Proc. Natl Acad. Sci. USA* **107**, 3087–3092 (2010).
- Katoh, Y. *et al.* Silencing the constitutive active transcription factor CREB by the LKB1-SIK signaling cascade. *FEBS J.* **273**, 2730–2748 (2006).
- Fu, A. & Screaton, R. A. Using kinomics to delineate signaling pathways: control of CRTC2/TORC2 by the AMPK family. *Cell Cycle* **7**, 3823–3828 (2008).
- Ahn, S. *et al.* A dominant negative inhibitor of CREB reveals that it is a general mediator stimulus-dependent transcription of c-fos. *Mol. Cell. Biol.* **18**, 967–977 (1998).
- Iourgenko, V. *et al.* Identification of a family of cAMP response element-binding protein coactivators by genome-scale functional analysis in mammalian cells. *Proc. Natl Acad. Sci. USA* **100**, 12147–12152 (2003).
- Conkright, M. D. *et al.* TORCs: transducers of regulated CREB activity. *Mol. Cell* **12**, 413–423 (2003).
- Lumeng, C. N., Deyoung, S. M., Bodzin, J. L. & Saltiel, A. R. Increased inflammatory properties of adipose tissue macrophages recruited during diet-induced obesity. *Diabetes* **56**, 16–23 (2007).
- Landsberg, L. Feast or famine: the sympathetic nervous system response to nutrient intake. *Cell. Mol. Neurobiol.* **26**, 497–508 (2006).
- Carmen, G. Y. & Victor, S. M. Signalling mechanisms regulating lipolysis. *Cell. Signal.* **18**, 401–408 (2006).
- Cannon, B. & Nedergaard, J. Brown adipose tissue: function and physiological significance. *Physiol. Rev.* **84**, 277–359 (2004).
- Seale, P., Kajimura, S. & Spiegelman, B. M. Transcriptional control of brown adipocyte development and physiological function—of mice and men. *Genes Dev.* **23**, 788–797 (2009).
- Cederberg, A. *et al.* FOXO2 is a winged helix gene that counteracts obesity, hypertriglyceridemia, and diet-induced insulin resistance. *Cell* **106**, 563–573 (2001).
- Qi, L. *et al.* Adipocyte CREB promotes insulin resistance in obesity. *Cell Metab.* **9**, 277–286 (2009).
- Wu, Z. *et al.* Transducer of regulated CREB-binding proteins (TORCs) induce PGC-1 $\alpha$  transcription and mitochondrial biogenesis in muscle cells. *Proc. Natl Acad. Sci. USA* **103**, 14379–14384 (2006).
- Roy, A. A. *et al.* RGS2 interacts with Gs and adenylyl cyclase in living cells. *Cell. Signal.* **18**, 336–348 (2006).
- Salim, S., Sinnarajah, S., Kehrl, J. H. & Dessauer, C. W. Identification of RGS2 and type V adenylyl cyclase interaction sites. *J. Biol. Chem.* **278**, 15842–15849 (2003).
- Sinnarajah, S. *et al.* RGS2 regulates signal transduction in olfactory neurons by attenuating activation of adenylyl cyclase III. *Nature* **409**, 1051–1055 (2001).
- Rohlf, E. M., Daniel, K. W., Premont, R. T., Kozak, L. P. & Collins, S. Regulation of the uncoupling protein gene (Ucp) by  $\beta$ 1,  $\beta$ 2, and  $\beta$ 3-adrenergic receptor subtypes in immortalized brown adipose cell lines. *J. Biol. Chem.* **270**, 10723–10732 (1995).
- Granneman, J. G. Expression of adenylyl cyclase subtypes in brown adipose tissue: neural regulation of type III. *Endocrinology* **136**, 2007–2012 (1995).
- Freson, K. *et al.* -391 C to G substitution in the regulator of G-protein signalling-2 promoter increases susceptibility to the metabolic syndrome in white European men: consistency between molecular and epidemiological studies. *J. Hypertens.* **25**, 117–125 (2007).
- Dentin, R., Hedrick, S., Xie, J., Yates, J. III & Montminy, M. Hepatic glucose sensing via the CREB coactivator CRTC2. *Science* **319**, 1402–1405 (2008).
- Qi, L. *et al.* TRB3 links the E3 ubiquitin ligase COP1 to lipid metabolism. *Science* **312**, 1763–1766 (2006).
- Goodarzi, M. O. *et al.* Lipoprotein lipase is a gene for insulin resistance in Mexican Americans. *Diabetes* **53**, 214–220 (2004).
- Goodarzi, M. O. *et al.* Determination and use of haplotypes: ethnic comparison and association of the lipoprotein lipase gene and coronary artery disease in Mexican-Americans. *Genet. Med.* **5**, 322–327 (2003).
- Bild, D. E. *et al.* Multi-ethnic study of atherosclerosis: objectives and design. *Am. J. Epidemiol.* **156**, 871–881 (2002).
- Livak, K. J. Allelic discrimination using fluorogenic probes and the 5' nuclease assay. *Genet. Anal.* **14**, 143–149 (1999).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This study was supported in part by National Institutes of Health grants R01-DK049777, R01-DK083834, P30-DK063491, R01-HL088457, R01-DK79888, R01-HL071205, N01-HC95159, N02-HL64278 and M01-RR00425 (General Clinical Research Center Grant from the National Center for Research Resources, The J.W. Kieckhefer Foundation, The Clayton Medical Research Foundation, Inc., The Helmsley Foundation, the Cedars-Sinai Winnick Clinical Scholars Award (to M.O.G.) and the Cedars-Sinai Board of Governor's Chair in Medical Genetics (J.I.R.). We thank B. Beutler and O. Siggs (The Scripps Research Institute) for help with macrophage studies.

**Author Contributions** Y.S., J.A. and M.M. conceived the project and experimental design. M.O.G. analysed human data. Y.S., J.A., H.I., R.B., J.K., J.G., M.I., J.P., M.F.H., P.K.S., N.G., L.V. and N.M. performed experiments and data analysis. X.G. designed and supervised human data analysis. J.C. performed statistical analysis. M.R.G. performed genotyping experiments. Y.-D.I.C. was involved in study design. K.D.T. designed and supervised human genotyping experiments. W.A.H. and J.I.R. designed and conceived the MACAD study. Y.S., M.O.G. and M.M. wrote the paper. All authors discussed the results and commented on the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to M.M. ([montminy@salk.edu](mailto:montminy@salk.edu)).

## METHODS

**Animal studies.** Mice were housed in a temperature-controlled environment under a 12 h light:dark cycle with free access to water and standard rodent chow diet (LabDiet 5001). For HFD studies, 6- to 8-week-old mice were transferred to a 60% HFD (Research Diets, D12492). Magnetic resonance imaging scans for fat and lean mass were performed using an Echo MRI-100 instrument according to the manufacturer's instructions. ISO ( $10 \text{ mg kg}^{-1}$ ) and leptin ( $3 \text{ mg kg}^{-1}$ ; once a day for 3 d) were administered intraperitoneally. All animal procedures were performed with an approved protocol from the Salk Institute Animal Care and Use Committee.

**Crt3<sup>-/-</sup> mice.** The targeting vector was constructed by replacing exon 1 of the *Crt3* gene, which encodes the CREB binding domain, with a phosphoglycerol kinase-neomycin selection cassette. The vector also contained a phosphoglycerol kinase-diphtheria toxin-A cassette for negative selection. The targeting vector was linearized and electroporated into R1 embryonic stem cells. G418-resistant clones were screened for homologous recombination by Southern blot analysis. *Crt3<sup>-/-</sup>* mice were backcrossed to C57/BL6 for up to three generations for metabolic studies.

**CRE-luciferase transgenic mice.** A cassette of eight tandem full CREB binding sites (CRE) in the context of a minimal CFTR promoter (−126 base pairs of 5' flanking sequence from the human *CFTR* gene) was amplified by PCR from lysed recombinant CRE-luc adenovirus<sup>17</sup>. The CRE-luciferase transgene with flanking H19 insulator sequences was microinjected into 129/Sv oocytes for implantation into pseudo-pregnant female mice.

Transgenic founders were identified by PCR analysis of genomic DNA. Founder lines were backcrossed onto albino C57/BL6 mice (C57BL/6-Tyrc-2J, Jackson) for three generations. To maintain consistent copy number, hemizygous transgenic mice were bred to wild-type mice; the transgene was segregated to 50% of offspring in each litter and was stable for at least three generations. *In vivo* luciferase activity was measured by intra-vital imaging. Mice were anaesthetized using isoflurane gas (2% to effect), injected with D-luciferin ( $150 \text{ mg kg}^{-1}$ , intra-peritoneally) and imaged in an IVIS-100 instrument (Caliper Life Sciences) for 1–5 min.

**Genotyping.** Genomic DNA was prepared from tail biopsies and genotyped using the following primer sets: CRTC3 wild-type allele, CCTGAGTTATTGGCGGATGT and CACTCAGGCTGTAGCAAGCA; CRTC3 knockout allele, ATGGAAGGATTGGAGCTACG and CACTCAGGCTGTAGCAAGCA; CRE-luciferase transgene, GCTGGGCGTTAATCAGAGAG and TTTTCCGTCATC GTCTTTCC.

**Histology.** Mouse tissues were fixed in zinc-buffered formalin (Anatech) and paraffin embedded. Sections ( $5 \mu\text{m}$ ) were used for haematoxylin and eosin staining or immunohistochemistry. For immunohistochemical staining of adipose tissue macrophages, rehydrated antigen retrieved sections were incubated with F4/80 (Serotec) antiserum and visualized by the avidin–biotin–complex method using the chromogen diaminobenzidine (Vector Labs).

**Immunoblot and chromatin immunoprecipitation.** Antibodies for CRTC3, phospho-HSL and phospho-PKA substrate were obtained from Cell Signaling. Immunoblots were performed as previously described<sup>16</sup>. Chromatin immunoprecipitation studies were performed as previously described<sup>15</sup>.

**Cell counting.** Images of sections ( $5 \mu\text{m}$ ) stained with haematoxylin and eosin were taken at  $\times 200$  magnification ( $1,300 \text{ pixels} \times 1,030 \text{ pixels}$  per picture). The National Institutes of Health ImageJ program was used to perform cell counts on brown adipose tissue sections.

**GTT, ITT.** For glucose tolerance testing, mice were fasted for 16 h and then injected with glucose ( $2 \text{ g kg}^{-1}$ , intraperitoneally). For insulin tolerance testing, mice were fasted 2–3 h and injected with insulin (Humulin;  $1 \text{ U kg}^{-1}$ , intraperitoneally).

**Metabolites.** Mouse blood was collected from the tail vein and glucose levels were measured with a One Touch Ultra Glucometer (Johnson & Johnson). Circulating insulin (Cayman), leptin (Millipore), adrenaline/noradrenaline (LDN GmbH & Co. KG) and free fatty acid (WAKO Chemicals) levels were determined by ELISA.

**Core body temperature.** Core body temperature was measured with a thermistor thermometer (Model 8402-10, Cole Parmer Thermometers).

**cAMP.** Tissue and cellular cAMP concentrations were determined using a cAMP ELISA kit according to the manufacturer's instruction (R&D Systems or Cayman Chemical Company). Cells were exposed to FSK ( $10 \mu\text{M}$ ) for times indicated.

**Indirect calorimetry.** Mice were individually housed for at least 3 days before calorimetry experiments. Food intake, locomotor activity, oxygen consumption and carbon dioxide production were simultaneously measured for individually housed mice with a LabMaster system (TSE Systems). Data were collected for 2–3 days and analysed. For leptin studies, mice were treated with saline or leptin ( $3 \mu\text{g g}^{-1}$ , intraperitoneally) 90 min before the onset of the dark cycle.

**Metabolic studies.** Rates of fatty-acid oxidation and lipolysis were measured as previously described<sup>31,40</sup>. Basal rates of glucose uptake were determined by measuring

the rate of [ $U\text{-}^{14}\text{C}$ ]2-deoxyglucose uptake by isolated soleus muscle. Pairs of soleus muscle were rapidly dissected from wild-type or *CRTC3<sup>-/-</sup>* mice. Isolated soleus strips were incubated twice for 15 min at  $37^\circ\text{C}$  in continuously gassed (95%  $\text{O}_2$ , 5%  $\text{CO}_2$ ) Krebs-Henseleit bicarbonate buffer containing 5.5 mM glucose and 0.5 mM oleate complexed to 2% fatty-acid-free bovine serum albumin (Millipore). After the 30 min pre-incubation period, muscle strips were transferred into fresh identical buffer with [ $U\text{-}^{14}\text{C}$ ]2-deoxyglucose ( $2 \mu\text{Ci ml}^{-1}$ ). Soleus strips were continuously gassed and incubated at  $37^\circ\text{C}$  for 30 min. At the end of the experiment, individual soleus strips were washed three times with ice-cold Hank's balanced salt solution. Muscle strips were weighed and incubated for 1 h at  $55^\circ\text{C}$  in digestion buffer (50 mM Tris, pH 8; 100 mM NaCl, 100 mM EDTA, 1% SDS and  $500 \mu\text{g ml}^{-1}$  proteinase K). All of the digested tissue was transferred into scintillant (EcoLite) and the radioactivity was measured by liquid scintillation counting.

**MEFs.** Mouse embryos were obtained from gravid female mice at embryonic days 13–14. Embryos were minced, trypsinized and washed with PBS. MEFs were plated in DMEM with 10% FBS, and 1% penicillin–streptomycin.

**Primary adipocyte and stromal vascular fraction cultures.** Primary adipocytes and the stromal vascular fraction were isolated from epididymal WAT and BAT, as described previously<sup>31,40</sup>. Primary adipocytes fractions were plated in DMEM containing 5.5 mM glucose, 2% fatty-acid-free bovine serum albumin and 1% penicillin–streptomycin. For stromal vascular fraction cells, pellets from adipocyte isolations were washed three times with HDB, and cultured in DMEM with 10% fetal bovine serum and 1% penicillin–streptomycin.

**RNA studies.** Total RNA was isolated by Trizol (Invitrogen) and RNeasy Mini Kit (Qiagen). Total RNA ( $1\text{--}2 \mu\text{g}$ ) was used for complementary DNA synthesis with Superscript II according to the manufacturer's instruction (Invitrogen). Relative mRNA amount was determined by real-time quantitative on a LightCycler 480 instrument (Roche).

**Statistics.** Data are presented as means  $\pm$  s.e.m. Statistical analysis was performed using an unpaired *t* test with GraphPad Prism software. Statistical significance is indicated as  $*P < 0.05$ ,  $**P < 0.01$  and  $***P < 0.001$ . All transient luciferase assays were performed on at least three independent occasions.

**Human subjects.** Associations with adiposity parameters (weight, BMI, waist circumference, hip circumference, waist/hip ratio) were first assessed in participants in the Cedars-Sinai/University of California, Los Angeles MACAD study, a study of Mexican-American families from Los Angeles<sup>41,42</sup>. In the present report, 206 two-generation Mexican-American families were included, comprising 779 subjects (adult offspring of probands with coronary artery disease and the spouses of those offspring) who underwent anthropometric measurements and genotyping. By design, the offspring were free of diabetes and clinically manifest cardiovascular disease, thus avoiding secondary changes in phenotype caused by overt disease. All studies were approved by Human Subjects Protection Institutional Review Boards at University of California, Los Angeles and Cedars-Sinai Medical Center.

Confirmatory studies were undertaken in Mexican-American subjects from the MESA. A detailed description of the MESA study design and methods has been published previously<sup>43</sup>. Briefly, 6,814 participants 45–84 years of age who identified themselves as white (2,748), black (1,930), Hispanic/Latino (1,496) or Chinese (806) were recruited from six US communities between 2000 and 2002. To obtain a replication cohort most similar to that of MACAD, we studied MESA Hispanics with exclusion of those recruited from the New York site, as the latter are mainly from the Caribbean and may thus have genetic differences from the Mexican-Americans of MACAD<sup>45</sup>. This resulted in a cohort of 987 MESA Mexican-Americans.

To determine whether the genetic associations observed in Mexican-Americans would also be seen in other ethnic groups, we also examined 2,527 non-Hispanic white subjects from MESA who had available anthropometric data. We then also accessed data from two large consortia of non-Hispanic whites, CHARGE ( $n = 31,373$ )<sup>46</sup> and GIANT ( $n = 32,504$ )<sup>47</sup>, both of which had conducted genome-wide association studies of BMI. These datasets did not overlap in subjects.

**Genotyping of human samples.** Genotyping of SNP rs8033595 (S72N) in MACAD was performed using TaqMan MGB technology as previously described<sup>42,44</sup>. The genotyping success rate was 98.3%. In the MESA Mexican-American and white cohorts, S72N was represented by a proxy SNP rs3862434 (A/G,  $r^2 = 1$  with S72N in the Mexican-American cohort of the phase III HapMap data, and  $r^2 = 1$  in the Caucasian European cohort of the phase II HapMap) that was directly genotyped (not imputed) in the genome-wide association study conducted in MESA. In CHARGE and GIANT, rs8033595 was either directly genotyped or imputed, depending on the genome-wide association study arrays used in the individual cohorts comprising each consortium.

**Human genetic association analysis.** The MACAD cohort is composed of small families and marrying-in spouses. Therefore the generalized estimating equation



(GEE1 (ref. 48)) approach was selected to use data from all phenotyped subjects. We evaluated association using this robust variance estimation approach to test hypothesized associations between phenotypes and genotypes while accounting for familial correlations present in the data. The PROC GENMOD procedure in SAS (version 9.0, SAS Institute) was used for the association analysis in which a sandwich estimator was used assuming exchangeable correlation. Family was taken as the cluster factor, i.e. members from the same family were assumed to be correlated. The adiposity traits were log transformed to approximate conditional normality and homogeneity of variance better. An additive genetic model was assumed in all the association analyses. Analyses used age and sex as covariates, unless otherwise specified. The same analytical techniques were used to assess association of rs3862434 with adiposity traits within MESA Mexican-Americans and non-Hispanic whites. We also conducted an analysis combining the MACAD and MESA Mexican-Americans, in which 72N and the minor allele of rs3862434

were considered the same. In CHARGE and GIANT, we conducted a weighted  $z$ -score-based meta-analysis of association of rs8033595 with BMI, combining  $P$  values obtained from a similar  $z$ -score-based meta-analysis conducted in each consortium. The meta-analysis was performed using the program METAL (<http://www.sph.umich.edu/csg/abecasis/metal/>).

45. Choudhry, S. *et al.* Population stratification confounds genetic association studies among Latinos. *Hum. Genet.* **118**, 652–664 (2006).
46. Heard-Costa, N. L. *et al.* NRXN3 is a novel locus for waist circumference: a genome-wide association study from the CHARGE Consortium. *PLoS Genet.* **5**, e1000539 (2009).
47. Willer, C. J. *et al.* Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nature Genet.* **41**, 25–34 (2009).
48. Zeger, S. L. & Liang, K. Y. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* **42**, 121–130 (1986).

# Origin of Saturn's rings and inner moons by mass removal from a lost Titan-sized satellite

Robin M. Canup<sup>1</sup>

The origin of Saturn's rings has not been adequately explained. The current rings are more than 90 to 95 per cent water ice<sup>1</sup>, which implies that initially they were almost pure ice because they are continually polluted by rocky meteoroids<sup>2</sup>. In contrast, a half-rock, half-ice mixture (similar to the composition of many of the satellites in the outer Solar System) would generally be expected. Previous ring origin theories invoke the collisional disruption of a small moon<sup>3,4</sup>, or the tidal disruption of a comet during a close passage by Saturn<sup>5</sup>. These models are improbable and/or struggle to account for basic properties of the rings, including their icy composition. Saturn has only one large satellite, Titan, whereas Jupiter has four large satellites; additional large satellites probably existed originally but were lost as they spiralled into Saturn<sup>6</sup>. Here I report numerical simulations of the tidal removal of mass from a differentiated, Titan-sized satellite as it migrates inward towards Saturn. Planetary tidal forces preferentially strip material from the satellite's outer icy layers, while its rocky core remains intact and is lost to collision with the planet. The result is a pure ice ring much more massive than Saturn's current rings. As the ring evolves, its mass decreases and icy moons are spawned from its outer edge<sup>7</sup> with estimated masses consistent with Saturn's ice-rich moons interior to and including Tethys.

The Jovian and Saturnian regular satellites are believed to have formed within circumplanetary disks of gas and solids produced during the end stages of nebular gas inflow to the planets<sup>6,8,9</sup>. Recent work predicts that multiple generations of Titan-sized satellites formed and were lost to collision with Saturn<sup>6</sup>. Each satellite grows no larger than a critical mass, at which point its orbit spirals into the planet owing to density wave interactions with the gas disk<sup>6</sup> (see Supplementary Information). This critical mass is comparable to that of Titan (mass  $M_T = 1.35 \times 10^{26}$  g and solid-body radius  $R_T = 2,575$  km)<sup>6</sup>. The overall process produces a final satellite system with either several large satellites (like Jupiter's Galilean satellites) or a single large satellite (like Saturn's Titan)<sup>6,9</sup>. A Saturn-like system (Table 1 and Fig. 1) can be produced when large, Titan-sized interior satellites spiral into the planet and are lost as gas inflow ends<sup>6</sup>.

A large satellite on an approximately circular orbit becomes unstable interior to a distance known as the classical Roche limit,  $a_R \equiv 2.456 R_S (\rho_S/\rho)^{1/3}$ , where  $R_S = 58,232$  km and  $\rho_S = 0.687$  g cm<sup>-3</sup> are Saturn's current mean radius and density, and  $\rho$  is the satellite's mean density. For a Titan-like mean density ( $\rho_T = 1.88$  g cm<sup>-3</sup>),  $a_R = 1.76 R_S$ . What happens once a large satellite drifts within  $a_R$  depends on its interior structure. An undifferentiated, uniform composition satellite disrupts completely. However, by the time a large ice-rock satellite approached the Roche limit, it would probably have undergone substantial ice melting and thus have a differentiated interior.

The energy associated with forming a Titan-sized satellite heats its interior to near the melting point for ice and may cause partial melting<sup>10</sup>. In addition, for even slightly non-circular orbits, the time-varying distortion of the satellite's shape by the planet heats the satellite's interior at a rate<sup>11</sup>  $dE/dt \approx (21/2) M_p \Omega^3 r^2 (R/r)^5 (k_2/Q) e^2$ , where  $M_p$  is the planet's mass,  $\Omega$  and  $r$  are the satellite's orbital frequency and orbital radius,  $R$  is

the satellite's radius,  $(k_2/Q)$  is the ratio of the satellite's Love number to its tidal dissipation factor (an uncertain quantity but plausibly within the range<sup>11</sup>  $10^{-3} < (k_2/Q) < 10^{-1}$ ; Supplementary Information), and  $e$  is the satellite's eccentricity. For example, with  $e = 0.002$  (Fig. 1 legend; Supplementary Information),  $r = 1.7 R_S$ ,  $(k_2/Q) = 10^{-2}$  and  $R = R_T$ , tidal heating dissipates about  $5 \times 10^9$  erg g<sup>-1</sup> in a Titan-mass satellite ( $M = M_T$ ) over an estimated orbital decay timescale of about  $10^4$  years (Supplementary Information), comparable to the latent heat of fusion of water ice ( $3 \times 10^9$  erg g<sup>-1</sup>). As ice melts, higher-density rock initially contained within the ice rapidly descends to the satellite's centre, so that melting in a satellite's outer layers creates an outer pure-ice mantle overlying a more rock-rich core<sup>10,12</sup>. For  $M = M_T$ , the separation of rock from ice becomes energetically self-sustaining<sup>13</sup> once about 50% of the rock has migrated to the satellite's centre<sup>12,13</sup>, and in this case the satellite differentiates into an ice mantle and a core that is pure rock or rock and metal.

When a differentiated satellite drifts within its Roche limit, tides first strip material from its outer, lower-density layers. The removal of low-density material causes the satellite's mean density to increase until the remnant satellite is marginally stable. As the satellite spirals inward, this process regulates  $\rho$  to approximately the local critical value for stability at the satellite's semi-major axis  $a$ ,  $\rho_{\text{crit}} \equiv \rho_S (2.456)^3 (R_S/a)^3$ , until the remnant satellite either collides with the planet or fully disrupts.

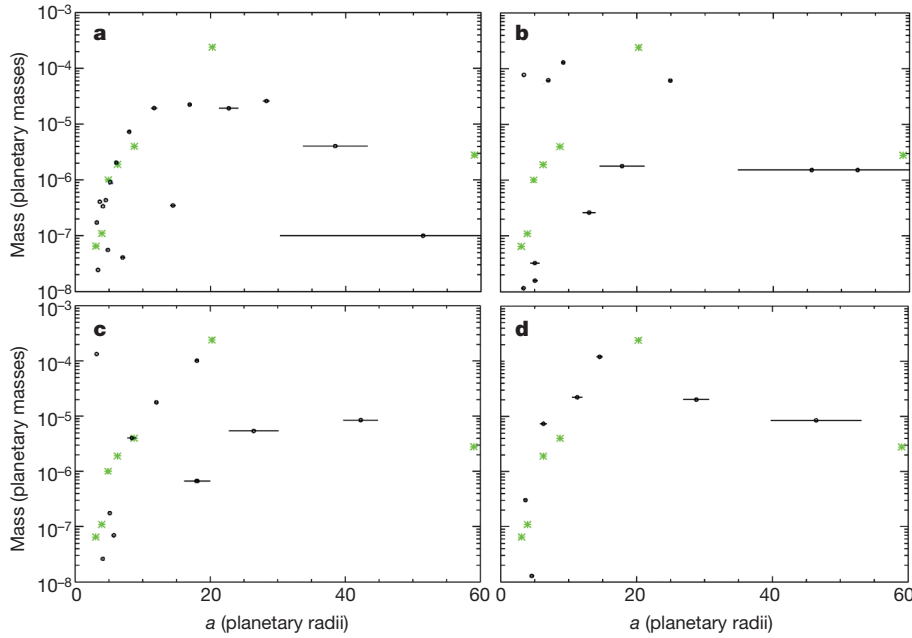
As a simple example, consider a two-layer satellite with an ice mantle overlying a core of uniform density  $\rho_{\text{core}}$ . The satellite's initial mean density  $\bar{\rho}_0$  determines the distance at which the tidal loss of ice begins [ $a_{\text{max}} \equiv a_R(\rho = \bar{\rho}_0)$ ], while  $\rho_{\text{core}}$  sets the distance at which the core disrupts [ $a_{\text{rock}} \equiv a_R(\rho = \rho_{\text{core}})$ ]. The satellite sheds only ice across the region:

**Table 1 | Saturn's rings and inner moons**

Object	$a/R_{\text{eq}}$	Mass ( $10^{22}$ g)	Mean radius (km)	Density (g cm <sup>-3</sup> )
B ring	1.5–1.9	2–10	–	–
A ring	2.0–2.3	0.6	–	–
Epimetheus	2.51	0.06	58.3	0.63
Janus	2.51	0.20	90.4	0.61
Mimas	3.09	3.75	198	1.15
Enceladus	3.95	10.8	252	1.61
Tethys	4.89	61.7	533	0.97
Dione	6.26	105	562	1.48
Rhea	8.74	230	764	1.23

Shown are key properties<sup>24,27</sup> of Saturn's most massive rings and inner moons with mean radii  $> 50$  km. Here orbital radii  $a$  are scaled by Saturn's equatorial radius according to convention ( $R_{\text{eq}} = 60,268$  km =  $1.035 R_S$ ). The total mass in the main rings, contained primarily in the B ring, is estimated to be<sup>24</sup> a few times  $10^{22}$  g to  $10^{23}$  g. Saturn's inner satellites interior to and including Tethys are, as a group, unusually ice-rich, with a mass-weighted average density  $\bar{\rho}_{\text{sat}} \approx 1.07$  g cm<sup>-3</sup>. Disruptive collisions<sup>4</sup> and/or endogenic activity in the case of Enceladus<sup>27</sup> could have removed ice relative to rock owing to ice's higher volatility, so that the current inner satellite compositions probably provide lower limits on their initial ice fraction. Moons orbiting exterior to Saturn's synchronous radius evolve outward owing to tidal interaction with the planet. This implies that the moons interior to and including Tethys could have all been interior to<sup>28</sup> about  $4 R_S$  when they formed, consistent with their having been spawned from the outer edge of the rings. The model here proposes that Tethys, Enceladus and Mimas (or their progenitors) were spawned from a primordial massive ring as it spread diffusively and delivered material to the region outside the Roche limit. A similar process appears to be ongoing today, with the smaller inner moons (interior to and including Janus) probably forming in the last  $10^6$  to  $10^7$  years as a result of recent spreading of the A ring<sup>7</sup>. More distant Dione and Rhea have been exterior to about  $6 R_S$  throughout their history<sup>29</sup>, and thus probably formed independently of the rings.

<sup>1</sup>Planetary Science Directorate, Southwest Research Institute, 1050 Walnut Street, Suite 300, Boulder, Colorado 80302, USA.



**Figure 1 | Results of a satellite accretion simulation<sup>6</sup> that produced a Saturn-like system of satellites.** The satellite disk is supplied by an inflow of gas and solid particles from solar orbit, the rate of which decays exponentially with a time constant  $\tau_{\text{in}}$  that is comparable to the solar nebula lifetime. Satellites undergo inward type I migration on a timescale proportional to  $(M\sigma_g)^{-1}$ , where  $M$  is satellite mass, and  $\sigma_g$  is the disk gas surface density, with the latter proportional to the inflow rate<sup>6,8</sup>. Black circles show the simulated satellites (with horizontal lines proportional to orbital eccentricities); Saturn's satellites are shown as green stars. **a**, Multiple satellites form as solid material flows into the disk ( $t = 0.2 \tau_{\text{in}}$ ). Once satellites grow to a mass  $M$  of a few times  $10^{-5}$  planetary masses, they begin to migrate inward. **b**, At  $t = 1.4 \tau_{\text{in}}$ , the system resembles Jupiter's Galilean system, with four similarly sized large satellites. **c**, The inner three large satellites are lost to collision with Saturn, with the last

large satellite lost at time  $t \approx 3 \tau_{\text{in}}$ , when the inflow rate has slowed substantially and the disk gas density has decreased to  $\sigma_g \approx 10 \text{ g cm}^{-2}$ . The last lost satellite acquires most of its mass at distances of 25 to 30 planetary radii, where disk temperatures were low enough for water ice, implying that the satellite would have a Titan-like composition with approximately 50% rock and 50% ice. **d**, The final system at  $t = 10 \tau_{\text{in}}$  has a single large Titan-like satellite at  $a \approx 15$  planetary radii. The satellite eccentricities in these simulations reflect a balance between mutual gravitational interactions and eccentricity damping by density waves, yielding an average final eccentricity for the large satellites of  $\langle e \rangle \approx 0.02$  (ref. 6). Inclusion of eccentricity damping by tides raised on the satellites by the planet is estimated to reduce this to  $\langle e \rangle$  of a few times  $10^{-3}$  (Supplementary Information).

$$a_{\text{rock}} < a < a_{\text{max}} \\ 1.5R_S \left( \frac{3.0 \text{ g cm}^{-3}}{\rho_{\text{core}}} \right)^{1/3} < a < 1.8R_S \left( \frac{1.9 \text{ g cm}^{-3}}{\bar{\rho}_0} \right)^{1/3} \quad (1)$$

For a rock or a rock-metal core,  $\rho_{\text{core}}$  would be about 3.0 to 3.5  $\text{g cm}^{-3}$  (the latter is the density of Jupiter's rocky satellite, Io), while  $\bar{\rho}_0 \approx \rho_T$  is expected.

Planet contraction models<sup>14,15</sup> predict that 2 to 5 Myr after the peak rate of its gas accretion, the young Saturn's radius  $R_p$  would have had a value between about  $1.6R_S$  and  $1.5R_S$ . Mean observed nebular lifetimes<sup>16</sup> are about 3 Myr, so that the end of gas inflow to Saturn and the loss of the final large satellites interior to Titan would probably occur in this timeframe. For  $a_{\text{rock}} < R_p < a_{\text{max}}$ , a satellite will collide with the planet before its rocky core disrupts, so that it tidally sheds only ice. Although each lost satellite may have produced tidal debris (depending on the state of the planet at that time), such material would generally be removed through collision with the planet as it was shepherded by subsequent satellites migrating inward<sup>17</sup> or perhaps driven to high eccentricities by not-too-distant large satellites<sup>18</sup>. However, tidal debris from the last large satellite to be lost from the Saturnian system could survive.

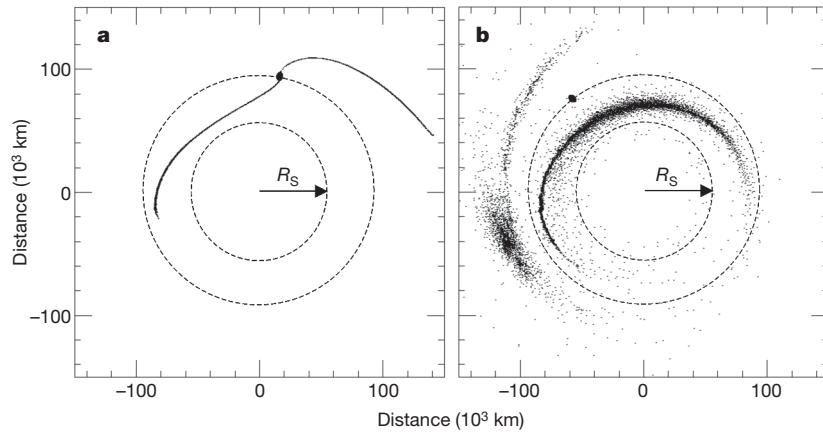
I use smooth particle hydrodynamics (SPH) to simulate tidal stripping from Titan-sized satellites (Fig. 2; Supplementary Information). The simulations predict initial tidal fragment orbital eccentricities of  $e_p \approx 0.1$ . Fragments would have radii of about 1 to 50 km, depending on the tensile strength of the satellite's outer ice shell<sup>5</sup> (Supplementary Information). Subsequent collisions between fragments occur with a characteristic velocity  $v \approx e_p r \Omega$ , where  $\Omega \equiv [GM_S/r^3]^{1/2}$  is the orbital frequency,  $M_S = 5.69 \times 10^{29} \text{ g}$  is Saturn's mass,  $G$  is the gravitational constant, and  $r\Omega$  is the orbital velocity at radius  $r$ , which is about

$20 \text{ km s}^{-1}$  at  $r = 1.7R_S$ . The collision energy per unit mass,  $v^2/2 \approx 10^{10} \text{ erg g}^{-1}$ , exceeds that needed to catastrophically disrupt<sup>19</sup> ice objects with radii of 1 km to 50 km ( $\sim 10^6 \text{ erg g}^{-1}$  to a few  $10^8 \text{ erg g}^{-1}$ ), so that collisions shatter the fragments into small particles.

Mutual collisions also rapidly drive the particles into a ring with nearly circular and co-planar orbits (Supplementary Information). To estimate the total mass of ice produced, I compute the equivalent circular orbit,  $a_{\text{eq}} \equiv a_p(1 - e_p^2)$ , for each SPH particle having the same angular momentum as its initial orbit with semi-major axis  $a_p$  and eccentricity  $e_p$ . I then compare  $a_{\text{eq}}$  to an estimate (Supplementary Information) of the ice stability distance  $a_{\text{ice}}$ , obtained by balancing heating of the disk by the planet's luminosity, occurring with a rate per unit area of the disk  $\dot{E}_p \approx (18/7)\sigma_{\text{S-B}}T_p^4(R_p/r)^2(c/r\Omega)$ , with radiative cooling from the disk surfaces, occurring with a rate per unit area of the disk  $\dot{E}_{\text{rad}} = 2\sigma_{\text{S-B}}T_d^4$ . Here  $\sigma_{\text{S-B}}$  is the Stefan-Boltzmann constant,  $T_p$  and  $R_p$  are the planet's temperature and radius determined by planet contraction models<sup>14,15</sup>,  $c \propto T_d^{1/2}$  is the speed of sound in the gas, and  $T_d$  is the disk temperature at radius  $r$ . Setting  $\dot{E}_p = \dot{E}_{\text{rad}}$  and solving for the distance at which  $T_d = 200 \text{ K}$  gives  $a_{\text{ice}}/R_S \approx 1.6(T_p/400 \text{ K})^{8/3}(R_p/1.5R_S)^{4/3}$ . As a Titan-like satellite orbit spirals towards the planet, up to a few times  $10^{25} \text{ g}$  of thermally stable ice particles is produced (Table 2).

The earliest dynamical evolution of the tidally stripped debris is dominated by gravitational interactions with the remnant satellite (Supplementary Information), but after the satellite collides with Saturn, the debris settles into a pure ice ring containing up to a few times  $10^{25} \text{ g}$ , with a surface density  $\sigma_{\text{ring}} \approx 10^5 \text{ g cm}^{-2}$ , and a ring outer edge at the Roche limit for ice:  $a_{\text{R,ice}} \approx 2.2R_S$ . The gas disk at this late time is estimated to have a much lower surface density than the ring, with  $\sigma_g \leq 10 \text{ g cm}^{-2}$  (Fig. 1; Supplementary Information). Describing





**Figure 2 | SPH simulation showing the tidal removal of ice from a differentiated, Titan-mass satellite.** SPH represents matter as particles, which are here evolved in response to gravity, shock dissipation, and pressure using the equation of state M-ANEOS<sup>29,30</sup> (Supplementary Information). Type I migration and tidal interaction with the planet cause the satellite's orbit to spiral inward from its initial Roche limit ( $a_{\max}$ ) to the planet's surface ( $R_p$ ) in about  $10^4$  years (Supplementary Information). The satellite's evolution across this region is tracked with a series of SPH simulations. The satellite starts with  $a = a_{\max}$  and is evolved for several orbits with SPH to simulate tidal mass removal and the establishment of a stable satellite remnant. The remnant satellite is then shifted inward by  $\Delta a \approx 10^{-2}a$  (Supplementary Information) and re-simulated, with the process repeated until  $a$  is small enough that the satellite's rocky core disrupts, which determines  $a_{\text{rock}}$ . Frames here show tidal stripping from a satellite with a composition of 55% ice and 45% silicate + metal

drag by the gas disk on the ring as a shear stress on the disk surfaces gives a ring decay timescale<sup>3,20</sup> of  $\tau_{\text{gd}} \approx 14\text{Re}(\sigma_{\text{ring}}/\sigma_g)(GM_S/c^3)$ , which is about  $10^7$  years for these surface densities,  $T_d \approx 200$  K and a Reynolds number  $\text{Re} \approx 10^2$  (Supplementary Information). This exceeds the expected persistence time of the gas disk (nominally  $\sim 10^6$  years), so the ring survives.

Interparticle collisions cause the ring to spread and decrease in mass as inward-flowing material is lost and outward-diffusing material accumulates into moons<sup>7,21</sup>. The spreading timescale for a massive ring is<sup>22,23</sup>:

$$\tau_v \approx \left( \frac{r\Omega^2}{G\sigma_{\text{ring}}} \right)^2 \frac{\Omega^{-1}}{C} \approx 10^5 \text{ years} \left( \frac{10}{C} \right) \left( \frac{10^5 \text{ g cm}^{-2}}{\sigma_{\text{ring}}} \right)^2 \left( \frac{1.8R_S}{r} \right)^{5/2} \quad (2)$$

where  $C(r)$  is a scaling factor<sup>23</sup> of  $\sim 10$  in the B ring and  $\sim 40$  at  $a_{R, \text{ice}}$ . The initially massive ring envisioned here would, over the age of the Solar System, have decreased in mass<sup>21</sup> such that  $\tau_v \approx 4.5$  Gyr, implying

**Table 2 | Results of SPH simulations of tidal stripping**

Satellite composition	$a_{\max}$	$R_{p, \min}$	$a_{\text{ice}}$	$M_{\text{ring}} (a_{\text{eq}} \geq a_{\text{ice}})$
50% ice, 15% hydrated silicate and 35% silicate	1.7	1.5	1.8	$2 \times 10^{25}$ g
50% ice and 50% silicate	1.7	1.5	1.8	$2 \times 10^{25}$ g
55% ice and 45% silicate + metal	1.7	1.4	1.5	$3 \times 10^{25}$ g
45% ice and 55% silicate + metal	1.6	1.4	1.5	$2 \times 10^{25}$ g

The first four columns list the compositions of the simulated satellites (which all had  $M = M_T$ ), the semi-major axis at which tidal loss of ice commences ( $a_{\max}$ ), the smallest planet radius consistent with pure ice loss ( $R_{p, \min} = a_{\text{rock}}$ ), and an estimate (see main text and Supplementary Information) of the ice stability radius in the circumplanetary disk ( $a_{\text{ice}}$ ), based on planet contraction models<sup>14,15</sup>. All distances are shown in units of Saturn's mean radius,  $R_S$ . For a given satellite composition, the mass of tidally stripped ice is maximized if the planet's surface at the time of the satellite's orbital decay is located at  $R_p = R_{p, \min}$ , so that the satellite is consumed by the planet just before the loss of rock from its core. A more extended planet with  $R_p > R_{p, \min}$  causes an earlier collision and a reduced mass of stripped ice, while a smaller planet radius and a later collision would lead to rocky fragments inconsistent with the composition of Saturn's rings. The final column shows the mass of ice particles having an equivalent circular orbit  $a_{\text{eq}}$ , exterior to  $a_{\text{ice}}$  assuming  $R_p = R_{p, \min}$ . Similar total ring masses can be estimated analytically (Supplementary Information). The  $a_{\text{eq}}$  distance is representative of the orbital radius to which the mass of an initial fragment will settle after undergoing dissipative collisions but before substantial diffusive spreading of the ring. By comparing  $a_{\text{eq}}$  to  $a_{\text{ice}}$ , I assume ring particles are in thermal equilibrium with the circumplanetary gas disk. If they instead radiate into the cooler background nebula, ice particles could be thermally stable interior to  $a_{\text{ice}}$  (Supplementary Information).

(see Table 2) orbiting a Saturn-mass planet at  $a = 0.97a_{\max}$  after 8 simulated hours (a) and 25 h (b). Distances are shown in units of  $10^3$  km; for comparison, Saturn's B and A rings lie between  $\sim 92,000$  and  $137,000$  km from the centre of Saturn. Dashed circles indicate the satellite's orbit and Saturn's current mean radius,  $R_S$ ; Saturn's radius at the time of the satellite's decay was probably  $R_p \approx 1.5R_S$  (refs 14 and 15; Supplementary Information). Material originating from the satellite's ice mantle is lost through its inner and outer Lagrange points (L1 and L2), leading to particles on highly eccentric orbits (with  $e \approx 10^{-1}$ ) with semi-major axes interior and exterior to that of the satellite, respectively (Supplementary Information). Subsequent collisions between particles will tend to circularize their orbits, and the clumps seen in b are transient features. Interior particles will probably collide directly with the planet or be driven into the planet by the satellite, while exterior particles can supply the ring.

that  $\sigma_{\text{ring}}$  would now be a few times  $10^2 \text{ g cm}^{-2}$ , consistent with current estimates for the B ring<sup>24</sup>.

Ring material spreading beyond the Roche limit accretes to form icy moons<sup>7</sup>. Each moon spawned from the ring's outer edge grows until it reaches a mass such that the timescale for its recoil from the ring due to resonant interactions,  $\tau_{\text{recoil}}$ , is comparable to the timescale for the ring's outward diffusion<sup>7</sup>. With<sup>17</sup>  $\tau_{\text{recoil}} \approx M_S^2[(a_m - r)/a_m]^3 / [1.68a_m^2\sigma_{\text{ring}}\Omega m_m]$ , setting  $\tau_{\text{recoil}} \approx \tau_v$  gives a characteristic moon mass:

$$m_m^* \sim r^2\sigma_{\text{ring}} \left( \frac{C}{1.68} \right) \left( \frac{r}{a_m} \right)^{1/2} \left( \frac{a_m - r}{a_m} \right)^3 \quad (3)$$

that depends on  $\sigma_{\text{ring}}$  and the position  $a_m$  at which the moon accretes relative to the ring edge  $r$ . Accretion models<sup>25</sup> find  $a_m/a_R \approx 1.1$  to  $1.2$ . With  $1.1 < (a_m/a_{R, \text{ice}}) < 1.2$ ,  $r = a_{R, \text{ice}}$ , a uniform surface density ring, and  $C = 40$ , the estimated moon mass in grams is  $3 \times 10^{23}(\sigma_{\text{ring}}/10^5 \text{ g cm}^{-2}) < m_m^* < 2 \times 10^{24}(\sigma_{\text{ring}}/10^5 \text{ g cm}^{-2})$  (in approximate agreement with Supplementary Fig. 4 of ref. 7). This is comparable to Tethys' mass ( $6 \times 10^{23}$  g) for the initial  $\sigma_{\text{ring}}$  values predicted here.

Upon reaching a mass of about  $m_m^*$ , resonant torques drive the moon away from the ring until its most distant strong resonances migrate out of the ring<sup>7,26</sup>. For the 2:1 inner Lindblad resonance, this occurs when  $a_m = 1.6r$ , or when  $a_m \approx 3.6R_S$  for  $r = a_{R, \text{ice}}$ . At this point the moon is exterior to Saturn's early synchronous radius (Supplementary Information), so that it continues to evolve outward because of tidal interaction with the planet. As each moon recoils away from the ring, a new moon is spawned from the ring's outer edge, with the moon masses decreasing with time as  $\sigma_{\text{ring}}$  decreases<sup>7</sup>. As moons evolve outward, any mixing with material originating from outside the rings would increase their rock content somewhat relative to that of the rings. The densities of the moons interior to and including Tethys imply that as a group they contain about 90% ice and 10% rock (Table 1; Supplementary Information).

A primordial ring must avoid contamination by impacts of silicate-rich micrometeoroids throughout its 4.5-billion-year lifetime in order to produce the  $>90$  to 95% water-ice ring observed today. Previous

work<sup>2</sup> suggests that Saturn's current rings would be polluted in only a few times 10<sup>8</sup> years. However, the rings could be primordial and still unpolluted if the impact rate was overestimated and/or the rings' mass was underestimated<sup>1</sup>. During its extended mission, the Cassini spacecraft will indirectly sample the impact rate, and will directly measure the rings' current total mass<sup>1</sup>. While prior ring origin theories (Supplementary Information) have envisioned an initial ring comparable in mass to the current rings, the model here implies an initial ring that is several orders of magnitude more massive. A massive early ring would be less vulnerable to pollution by rock-rich impacts, and also has the advantage of providing sufficient mass and angular momentum (Supplementary Information) to account ultimately for both the current rings and the inner ice-rich Saturnian satellites.

Received 24 May; accepted 5 November 2010.

Published online 12 December 2010.

- Cuzzi, J. N. *et al.* An evolving view of Saturn's dynamic rings. *Science* **327**, 1470–1475 (2010).
- Cuzzi, J. N. & Estrada, P. R. Compositional evolution of Saturn's rings due to meteoroid bombardment. *Icarus* **132**, 1–35 (1998).
- Harris, A. W. in *Planetary Rings* (eds Greenberg, R. & Brahic, A.) 641–659 (University of Arizona Press, 1984).
- Charnoz, S., Morbidelli, A., Dones, L. & Salmon, J. Did Saturn's rings form during the Late Heavy Bombardment? *Icarus* **199**, 413–428 (2009).
- Dones, L. A recent cometary origin for Saturn's rings? *Icarus* **92**, 194–203 (1991).
- Canup, R. M. & Ward, W. R. A common mass scaling for satellites of gaseous planets. *Nature* **441**, 834–839 (2006).
- Charnoz, S., Salmon, J. & Crida, A. The recent formation of Saturn's moonlets from viscous spreading of the main rings. *Nature* **465**, 752–754 (2010).
- Canup, R. M. & Ward, W. R. Formation of the Galilean satellites: conditions of accretion. *Astron. J.* **124**, 3404–3423 (2002).
- Sasaki, T., Stewart, G. R. & Ida, S. Origin of the different architectures of the Jovian and Saturnian satellite systems. *Astrophys. J.* **714**, 1052–1064 (2010).
- Barr, A. C., Citron, R. I. & Canup, R. M. Origin of a partially differentiated Titan. *Icarus* **209**, 858–862 (2010).
- Showman, A. P., Stevenson, D. J. & Malhotra, R. Coupled orbital and thermal evolution of Ganymede. *Icarus* **129**, 367–383 (1997).
- Barr, A. C. & Canup, R. M. Origin of the Ganymede–Callisto dichotomy by impacts during the late heavy bombardment. *Nature Geosci.* **3**, 164–167 (2010).
- Friedson, A. J. & Stevenson, D. J. Viscosity of rock–ice mixtures and applications to the evolution of icy satellites. *Icarus* **56**, 1–14 (1983).
- Fortney, J. J., Marley, M. S. & Barnes, J. W. Planetary radii across five orders of magnitude in mass and stellar insolation: application to transits. *Astrophys. J.* **659**, 1661–1672 (2007).
- Marley, M. S., Fortney, J. J., Hubickyj, O. & Lissauer, J. J. On the luminosity of young Jupiters. *Astrophys. J.* **655**, 541–549 (2007).
- Haisch, K. E., Lada, E. A. Jr & Lada, C. J. Disk frequencies and lifetimes in young clusters. *Astrophys. J.* **553**, L153–L156 (2001).
- Goldreich, P. & Tremaine, S. The dynamics of planetary rings. *Annu. Rev. Astron. Astrophys.* **20**, 249–283 (1982).
- Goldreich, P. & Tremaine, S. The origin of the eccentricities of the rings of Uranus. *Astrophys. J.* **243**, 1062–1075 (1981).
- Benz, W. & Asphaug, E. Catastrophic disruptions revisited. *Icarus* **142**, 5–20 (1999).
- Goldreich, P. & Ward, W. R. The formation of planetesimals. *Astrophys. J.* **183**, 1051–1061 (1973).
- Salmon, J., Charnoz, S., Crida, A. & Brahic, A. Long-term and viscous evolution of dense planetary rings. *Icarus* **209**, 771–785 (2010).
- Ward, W. R. & Cameron, A. G. W. Disc evolution within the Roche limit. *Proc. Lunar Planet. Sci. Conf.* **IX**, 1205 (1978).
- Daisaka, H., Tanaka, H. & Ida, S. Viscosity in a dense planetary ring with self-gravitating particles. *Icarus* **154**, 296–312 (2001).
- Robbins, S. J., Stewart, G. R., Lewis, M. C., Colwell, J. E. & Sremcevic, M. Estimating the masses of Saturn's A and B rings from high-optical depth *N*-body simulations and stellar occultations. *Icarus* **206**, 431–445 (2010).
- Ida, S., Canup, R. M. & Stewart, G. R. Lunar accretion from an impact-generated disk. *Nature* **389**, 353–357 (1997).
- Canup, R. M. & Ward, W. R. A hybrid fluid/*N*-body model for lunar accretion. *Proc. Lunar Planet. Sci. Conf.* **XXXI**, 1916 (2000).
- Jaumann, R. *et al.* in *Saturn from Cassini–Huygens* (eds Dougherty, M. K., Esposito, L. W. & Krimigis, S. M.) 637–681 (Springer, 2009).
- Meyer, J. & Wisdom, J. Tidal heating in Enceladus. *Icarus* **188**, 535–539 (2007).
- Melosh, H. J. A hydrocode equation of state for SiO<sub>2</sub>. *Meteorit. Planet. Sci.* **42**, 2079–2098 (2007).
- Canup, R. M. Simulations of a late lunar-forming impact. *Icarus* **168**, 433–456 (2004).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** I thank W. R. Ward and L. Dones for detailed comments. Support from NASA's Outer Planets Research Program is gratefully acknowledged.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The author declares no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to the author ([robin@boulder.swri.edu](mailto:robin@boulder.swri.edu)).

# Pleats in crystals on curved surfaces

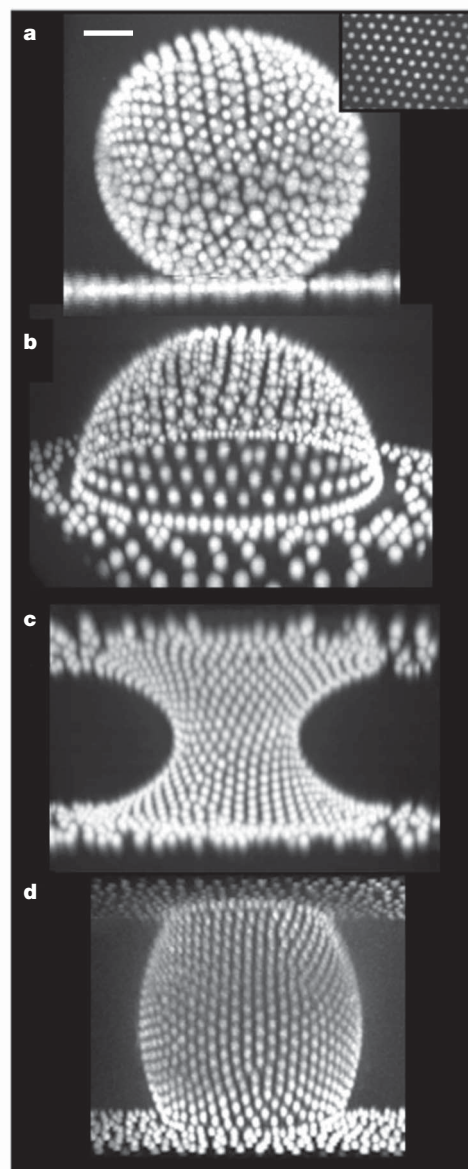
William T. M. Irvine<sup>1†</sup>, Vincenzo Vitelli<sup>2</sup> & Paul M. Chaikin<sup>1</sup>

Hexagons can easily tile a flat surface, but not a curved one. Introducing heptagons and pentagons (defects with topological charge) makes it easier to tile curved surfaces; for example, soccer balls based on the geodesic domes<sup>1</sup> of Buckminster Fuller have exactly 12 pentagons (positive charges). Interacting particles that invariably form hexagonal crystals on a plane exhibit fascinating scarred defect patterns on a sphere<sup>2–4</sup>. Here we show that, for more general curved surfaces, curvature may be relaxed by pleats: uncharged lines of dislocations (topological dipoles) that vanish on the surface and play the same role as fabric pleats. We experimentally investigate crystal order on surfaces with spatially varying positive and negative curvature. On cylindrical capillary bridges, stretched to produce negative curvature, we observe a sequence of transitions—consistent with our energetic calculations—from no defects to isolated dislocations, which subsequently proliferate and organize into pleats; finally, scars and isolated heptagons (previously unseen) appear. This fine control of crystal order with curvature will enable explorations of general theories of defects in curved spaces<sup>5–11</sup>. From a practical viewpoint, it may be possible to engineer structures with curvature (such as waisted nanotubes and vaulted architecture) and to develop novel methods for soft lithography<sup>12</sup> and directed self-assembly<sup>13</sup>.

Topological defects have played a crucial role in understanding the order, rigidity and melting of crystals and other phases of matter in two-dimensional flat space<sup>14,15</sup>. On a curved surface (Fig. 1), these particle-like excitations acquire a new life: they interact not only with each other, but with the curvature of the substrate. In a hexagonal lattice in which every particle has six nearest neighbours (Fig. 1, inset), there are two types of topological defects (Fig. 2): disclinations that disrupt orientational order and appear as points of local five-fold or seven-fold symmetry, (pentagons or heptagons, having topological charge  $\pm(2\pi/6)$ ), and dislocations, which disrupt translational order and appear as disclination dipoles ( $+/-$  pairs). That disclinations couple to curvature can be understood intuitively by taking a piece of paper, and adding, or removing, a  $\pi/3$  wedge to ‘make’ a disclination, Fig. 2c, d.

A host of new discoveries<sup>2,4,7,16</sup> have resulted from studies of these defects on the simplest curved surface: the sphere. With increasing size, the familiar 12-pentagon soccer ball pattern gives way to ‘scars’, pentagons dressed by strings of dislocations<sup>2–4</sup>. In this Letter, we introduce a different configuration of dislocations, namely, ‘pleats’—topologically uncharged grain boundaries with variable spacing that vanish on the surface. We experimentally investigate their interaction with curvature and show when pleats are energetically favoured over undefected crystals or topologically charged disclinations. Apart from experiments on spheres, and bubble bearing paraboloids<sup>17</sup>, the interaction of defects with variable curvature, negative curvature and surfaces of different topologies has remained largely unexplored experimentally and is of growing theoretical interest<sup>8–11,18–20</sup>.

The topology of surfaces<sup>21</sup> places a constraint on the total defect charge—for example the net charge on a sphere must be  $4\pi$ , as exemplified by a soccer ball that has 12 ( $=4\pi/(\pi/3)$ ) pentagons dispersed among hexagons. A hemisphere, or disk, requires half the topological



**Figure 1 | Colloidal crystals on curved oil-glycerol interfaces.**

**a–d**, Fluorescent PMMA particles bound, by image attraction, to oil-glycerol interfaces in the shape of spheres (**a**), domes (**b**), waists (**c**) and barrels (**d**) (see Supplementary Information section 2 for details of shape). The particles interact via a repulsive screened Coulomb interaction and, on a flat surface, arrange into a hexagonal crystal lattice (**a**, inset). The oil phase is a mixture of cyclohexyl bromide and dodecane that matches the refractive index of glycerol, allowing us to image particles on highly curved interfaces by confocal microscopy. Because of their topology, crystals on spheres and domes require a net defect charge of  $12 \times (2\pi/6)$  and  $6 \times (2\pi/6)$ , respectively, whereas waists and barrels require none. While spheres (**a**) have no boundary, the remaining surfaces (**b–d**) do, allowing topological defects a choice between boundary and bulk.

<sup>1</sup>Center for Soft Matter Research, Department of Physics, New York University, 4 Washington Place, New York, New York 10003, USA. <sup>2</sup>Instituut-Lorentz for Theoretical Physics, Leiden University, 271, Niels Bohrweg 2 Leiden, NL 2333 CA, The Netherlands. <sup>†</sup>Present address: James Franck Institute and Department of Physics, The University of Chicago, Chicago, Illinois 60637, USA.



charge of a sphere, whereas a cylinder can be defect-free with a requirement of 0 total charge.

Topology constrains the total charge, but it is energetics that determines the number and the arrangement of each charge. In elasticity theory<sup>15</sup>, disclinations appear as discrete interacting charges and the Gaussian curvature as a charge density, both acting as sources of stress:

$$\frac{1}{Y} \nabla^4 \chi = G(x) - \sum_{\alpha} q_{\alpha} \delta(x - x_{\alpha}) \quad (1)$$

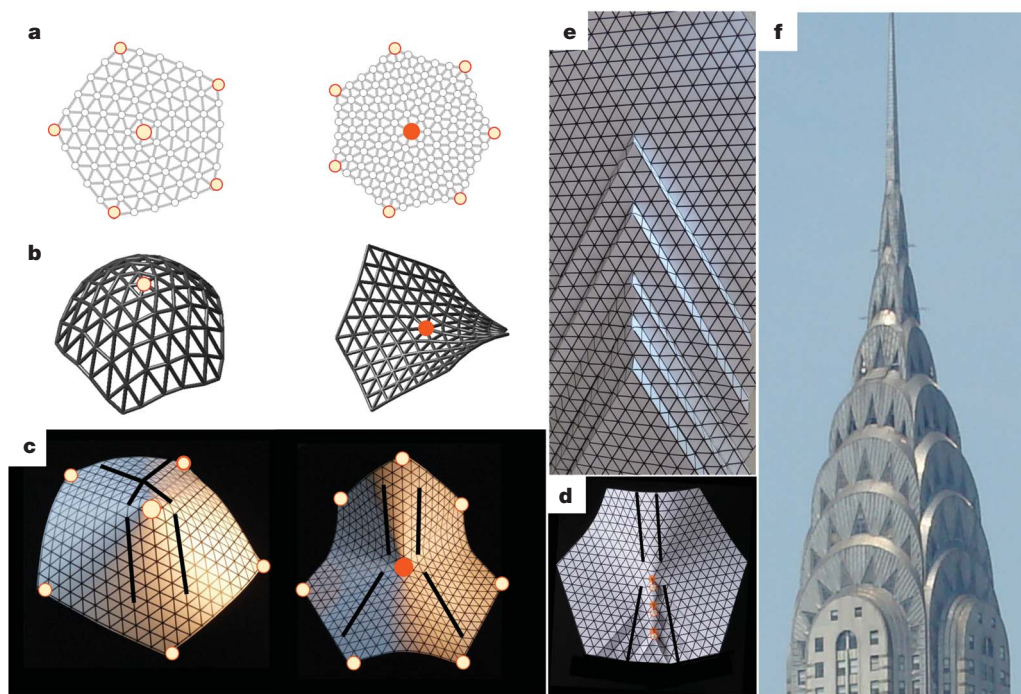
where  $G = \frac{1}{R_1 R_2}$  is the Gaussian curvature (with  $R_1$  and  $R_2$  the principal radii of curvature),  $q_{\alpha}$  is the charge of the disclination at  $x_{\alpha}$ ,  $\delta$  represents a Dirac delta function,  $Y$  is Young's modulus and  $\chi$  is the Airy stress function—and we will refer to the partial or complete cancellation of the effects of curvature and stress by topological charges (r.h.s. of equation (1)) as screening. For curvature concentrated at a point, as in our model paper disclinations of Fig. 2c, the screening can be perfect, the lattice is stress free and the energy  $E = \frac{1}{2Y} \int (\nabla^2 \chi)^2 dA = 0$  (where we have assumed no bending energy). For smooth surfaces, the screening is more subtle. Geometry provides some insight: consider a geodesic triangle drawn on a curved surface, for example, a sphere. Curvature causes lines to diverge or converge, affecting the angles at the vertices: the sum of the external angles will differ from  $2\pi$  by  $\Delta\theta = \int G dA$ . The same applies to any closed loop formed by connecting lattice sites and serves as a measure of angular strain. If instead, a disclination is encircled by the loop, by definition this adds/removes

a contribution of  $\pm(2\pi/6)$  regardless of the size of the loop. If sufficient curvature is enclosed, the angular stress generated by the disclination is screened on the outside. Within the patch the screening is incomplete and leads to an energetic cost. For later use we define  $\Omega = \frac{1}{\pi/3} \int G dA$ , the integrated curvature in units of disclinations.

The crystals we create consist of poly(methyl methacrylate) (PMMA) particles ( $\sim 2 \mu\text{m}$  diameter) that are bound to an oil–glycerol interface and repel each other. By index-matching the oil to the glycerol, we can image the full surfaces (Fig. 1, Methods). The first surfaces we investigate are domes (truncated spheres), created by droplets sitting on a coverslip, with a circular contact line. The contact angle, controlled by treating the glass surface, determines the solid angle of the spherical droplet.

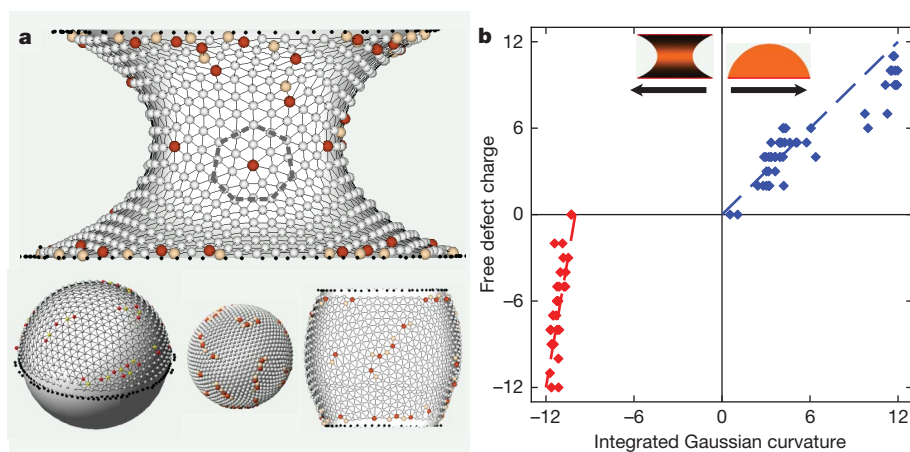
The domes (Fig. 3) exhibit both disclinations and scars as previously seen on spherical surfaces. However, for domes, the net disclination charge on the surface can vary as the dome inflates from a disk to a full sphere. In Fig. 3b we show the topological charge that is on the surface of the dome and detached from the boundary as a function of  $\Omega$ . We find the intuitive result that the detached charge varies approximately linearly with  $\Omega$ . That is, for a full sphere there are 12 pentagons ( $+\pi/3$ )s, for a hemisphere there are 6 pentagons ( $+\pi/3$ )s and on smaller fractions the two remain approximately proportional. Note that the topological requirement of a total of  $6 (+\pi/3)$ s is satisfied at all times by compensating charges on the boundary.

Negative curvature surfaces lack the simplicity and familiarity that we associate with positive curvature surfaces such as the sphere. For



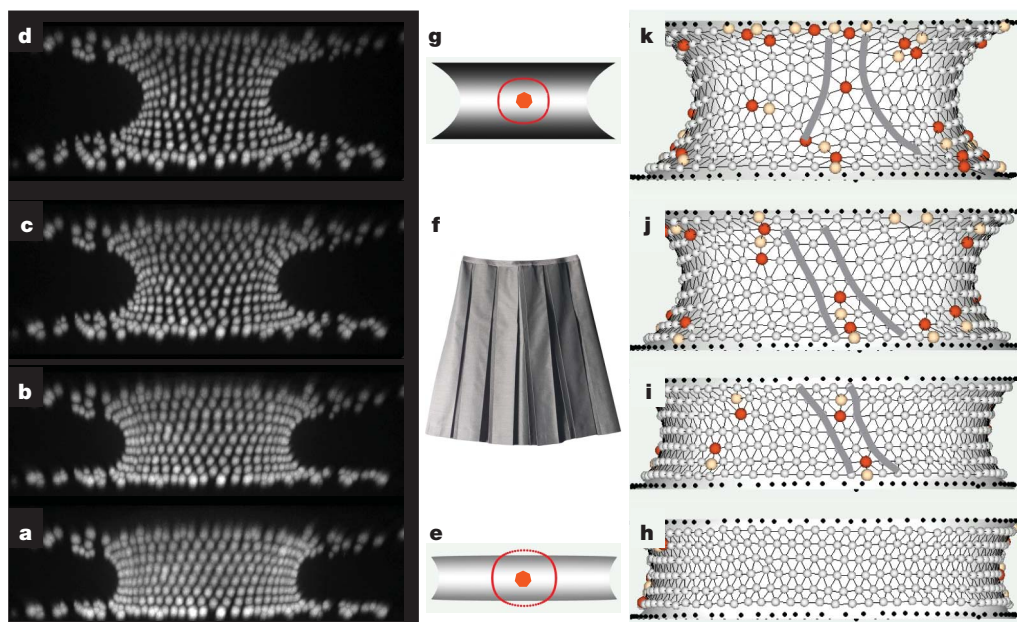
**Figure 2 | Disclinations and pleats in a hexagonal lattice.** **a**, Disclinations in a hexagonal lattice are topological defects that result from an extra (right panels) or missing (left panels)  $60^\circ$  crystalline wedge that matches their topological charge of  $+/- (2\pi/6)$  marked here by cream/brick circles, respectively. At the ‘core’ of the disclination is a corresponding five-fold (or seven-fold) coordinated particle. In flat space, disclinations produce a large amount of stress in a crystal, disrupting orientational order. **b**, This stress can be relieved by buckling the crystal to create a curved surface. Five-fold disclinations are sources of positive Gaussian curvature, seven-fold disclinations of negative curvature. **c**, This coupling can be intuitively understood by making disclinations out of paper (see Supplementary Information for instructions and cut-outs). Paper can be bent much more easily than it can be stretched/compressed, so it bends, resulting in a surface free of stress, with all the Gaussian curvature concentrated at the locations of the disclinations. Neighbouring crystal planes diverge on these surfaces, matching the geodesics of the curved surface. **d**, Dislocations,

uncharged pairs of seven- and five-fold disclinations, can also be made by folding and gluing hexagonal paper. A set of three closely spaced, aligned dislocations (7-5,7-5,7-5) are shown on an approximately relaxed sheet. This is a grain boundary which vanishes at the centre of the sheet—a ‘pleat’. Note that negative curvature emanates from the vanishing point of the pleat, as evidenced both by the buckling of the sheet and by the  $30^\circ$  divergence of parallel lines impinging from the top. **e**, A stress-free pleat can be achieved by allowing steps out of the surface. The pleat retains the property that width is added along the pleat length in proportion to the linear density of dislocations. **f**, The top of the Chrysler building in New York consists of four vertical pleats on a cylinder. Here the pleats are formed from dislocations in a square field. Counting from the top, steps 3, 4, 5, 6 are approximately equally spaced at 8.4 m and form a cone with no Gaussian curvature. A gradient in linear dislocation density is achieved by spacing the second and first step at 9.4 m and 9.8 m spacing, resulting in a spike with negative Gaussian curvature crowning the cone.



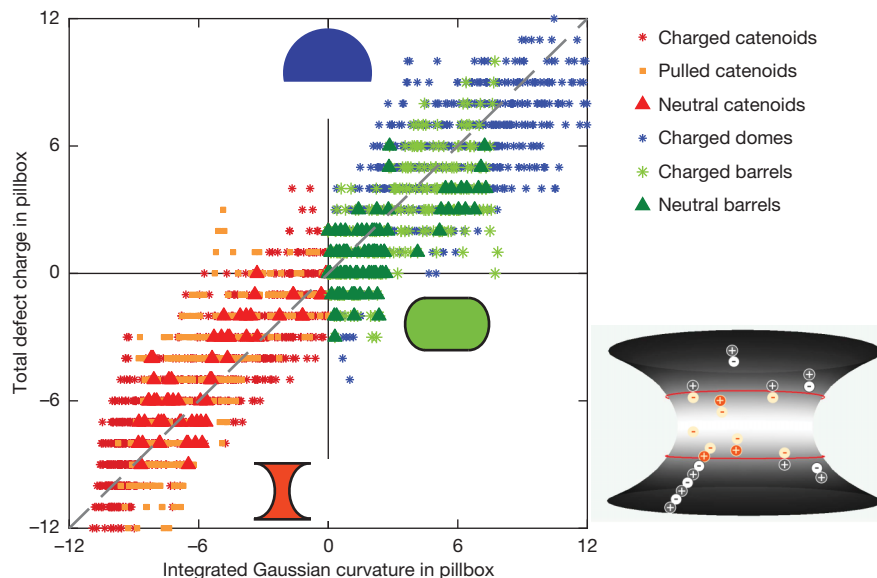
**Figure 3 | Topological charge on domes and waists.** **a**, Isolated seven-fold disclinations, observed here for the first time, can be seen at the neck of the capillary bridge (top). On the dome, sphere and barrel (bottom), the defects group into grain boundaries each having a single five-fold disclination in excess. The angular length of these scars is uniform, the number of dislocations depends only on  $R/a$ . The defect configurations on both surfaces are similar to the arrangement of defects on the surface of a sphere with no boundary. Both surfaces have uniform positive curvature. The black dots represent particles stuck to the glass surface, defects with charge  $+\pi/3$  are coloured brick, defects with charge  $-\pi/3$  are coloured cream. **b**, Detached topological charge on

domes and waists versus the integrated Gaussian curvature,  $\Omega$ . Detached charge is the sum of the charge of isolated disclinations and of scars that are not connected to the boundary. On domes (blue symbols), we find the intuitive result that there is approximately one ( $\pi/3$ ) topological charge per unit's  $\pi/3$  worth of integrated curvature. On negatively curved capillary bridges (red symbols), we find no net disclinations on the surface for an integrated curvature down to  $-10$ . For total curvature beyond the threshold of  $-10$ , disclinations rapidly fill the surface until  $12$  ( $-\pi/3$ ) disclinations match the  $\Omega = -12$  curvature. The dashed lines are guides to the eye.



**Figure 4 | Pleating and disclination unbinding on a stretched capillary bridge.** **a–d**, Confocal images of a capillary bridge that was stretched from quasi-cylindrical to highly curved. **h–k**, The corresponding reconstructed and triangulated surfaces. **a, h**, The compressed bridge exhibits very few defects. In this regime, the bridge is weakly curved and therefore a large patch of integrated curvature is required to screen the charge of a disclination. The circle that encloses such a patch is shown in **e**; clearly its radius exceeds the height of the bridge. As the bridge is stretched (**i–j**), we observe the appearance of dislocations, neutral disclination dipoles, polarized with the seven-fold defects towards the maximally negatively curved neck. The crystal planes can be seen to

diverge around these dislocation pairs, healing the stress induced by the negative curvature. As the bridge is stretched further, these dislocations proliferate, forming 'pleats', neutral grain boundaries that vanish in proximity of the neck. Uniform pleats on a skirt (**f**) reduce the circumference with height, producing a conical shape. Decreasing the pleat density with height would lead to negative curvature flaring the skirt. The divergence of the crystal planes at the opening of a high-density pleat (grey lines) is roughly half that observed around an isolated disclination (**k**). Finally, as the size of a patch required to screen the topological charge of an isolated seven-fold defect fits on the surface (**g**), we observe the appearance of isolated disclinations.



**Figure 5 | Disclination and polarization charge on the surfaces of waists, domes and barrels.** On each surface in our ensemble, we considered a patch ('pillbox') of variable size, for example as represented by the area between the red lines in the inset (bottom right) and computed the integrated Gaussian curvature (horizontal axis) versus the accumulated disclination plus polarization charge in the patch (vertical axis). All topological charges in the

example, it is not possible to embed a complete pseudo-sphere (sphere of negative curvature) in three-dimensional space. We create negative curvature surfaces by forming capillary bridges between two surfaces treated to have partially wetting glycerol contacts. These surfaces, shown in Figs 1, 3 and 4, are constrained to have constant mean curvature,  $\frac{1}{R_1} + \frac{1}{R_2}$ , and enclose a fixed volume. They are unduloids, nodoids and for some specific constructions catenoids or sections of spheres<sup>22,23</sup> (See Supplementary Information section 2). Topologically equivalent to cylinders and therefore requiring zero net charge, these surfaces typically have varying Gaussian curvature, as shown by the colour shading in Fig. 3b. On highly curved surfaces, we observe isolated  $-\pi/3$  disclinations (heptagons) for the first time (Fig. 3a). These are the negative curvature counterpart of the pentagons on the sphere and can also appear as 'scars' with an excess heptagon. However, if we plot (Fig. 3) the charge detached from the boundary versus  $\Omega$  for an ensemble of capillary bridges, we find no net topological charge on the surface for an integrated curvature down to  $-10$ . For total curvature beyond this threshold, disclinations rapidly fill the surface until 12 heptagons, each ( $-\pi/3$ ), match the  $\Omega = -12$  curvature. If no isolated charges are present below threshold, how does the crystal screen the curvature? A glance at surfaces below threshold (Fig. 4i, j) immediately suggests an answer: charge neutral dislocations.

To investigate this regime, we pulled apart the coverslips that confine the capillary bridges. As the boundary of the bridges is pinned and their volume conserved, their shape changes from almost cylindrical to highly curved, allowing us to follow the introduction of defects below threshold (see Fig. 4). The first defects that we observe as the (negative) curvature is increased are dislocations, dominantly polarized with their ( $-\pi/3$ ) disclinations pointing towards the region of highest negative curvature, the waist of the capillary bridge. As the integrated negative curvature is increased, these polarized dislocations proliferate and organize into lines. In flat space these lines, known as grain boundaries, close on themselves or span the whole sample. Here we observe neutral grain boundaries which vanish on the surface. They are analogous to fabric 'pleats'. In clothing, the different circumference of the waist and the hips is often accommodated by vertical pleats (Fig. 4f). Along the length of the pleat, extra fabric is added to the width.

patch (excluding the edges) contribute, whether isolated or connected by a pleat to the edge. Once pleats and scars are taken into account, we regain the intuitive result of a linear relation between Gaussian curvature and defect charge. Note that the triangular symbols correspond to samples with no detached charge on the surface.

At the point where the pleat begins, the opening angle  $\Delta\theta$  is determined by the dislocation line density  $n_d$  according to<sup>24</sup>

$$\Delta\theta = n_d 2a \frac{\pi}{6} \quad (2)$$

where  $a$  is the lattice spacing. If  $\Delta\theta$  perfectly matches the integrated Gaussian curvature, the stress induced by the substrate is completely screened. The smallest  $\Delta\theta$  is achieved by a pleat for which  $n_d$  is so low that only a single dislocation is present on a surface of characteristic length  $H$ . Equation (2) then gives a geometric criterion for the onset of pleating:  $\Delta\theta = |GdA| \approx a/H$  which can be arbitrarily small. Perfect screening cannot be achieved on a smooth surface. However, balancing the energy of a stretched but un-defected crystal,  $Y\Delta\theta^2 H^2$ , with the energy of a pleat,  $YaH\Delta\theta\log(\Delta\theta)$ , gives an energetic criterion for the onset of pleating,  $|GdA| \approx (a/H)\log(H/a)$ , that agrees with the geometric criterion up to logarithmic corrections (Supplementary Information).

There is a maximum opening angle for pleats  $\Delta\theta_{\max} = \frac{\pi}{6}$  that corresponds to the maximum dislocation line density  $n_d = \frac{1}{2a}$  in equation (2), see Fig. 2d. If we place these pleats along the axis of a cylinder the result is a cone, as illustrated in Fig. 2f, which shows the conical shape of the top of the Chrysler building, in New York. If the dislocation density in the pleats (or the density of pleats) varies, then we can have Gaussian curvature. On the Chrysler building, such a gradient in spacing produces a negative curvature spike that crowns the cone.

When our capillary bridges are strongly curved, multiple pleats and scars act in concert to relieve the strain induced by the curvature. This cooperative screening can be understood by treating dislocations as dipoles of disclinations. Just as in electrostatics, a divergence in polarization produces a polarization charge, pleat gradients can produce an effective disclination polarization charge, which along with isolated disclinations contributes to screen curvature so that

$$\int G dA \approx \sum_{\alpha=1}^n q_{\alpha} + \int dA \nabla \cdot P(x) \quad (3)$$

where  $P(x)$  is the dislocation density per unit area. In Fig. 5, we verify that the integrated Gaussian curvature over a 'patch' of variable size on each surface in our ensemble (including domes, barrels and waists) is approximately equal to the accumulated disclination and polarization



charges in the patch (excluding surface edges), for all regimes of deformation.

Pleats allow a finer screening of curvature than scars, as the angular deficit produced by disclinations is not tunable but quantized in units of  $\pi/3$ . Upon equating  $\pi/3$  to the Gaussian curvature multiplied by the area of a patch of radius  $r$ , we have a heuristic criterion for scarring  $r = \sqrt{\frac{R_1 R_2}{3}}$  that is consistent with our observations. If the distance to the boundary of our curved substrates is greater than  $r$ , isolated disclinations or scars are observed; if less than  $r$ , pleats arise as the generic mode of stress relaxation.

In summary, pleats (neutral dislocation lines that vanish on the on the surface) afford a much finer screening of Gaussian curvature than do disclinations of quantized topological charge, because the area surrounding the vanishing point of a pleat can have a continuously controllable angle deficit between 0 and  $\pi/6$ . Whether pleats may have a role to play in the design of geodetic-dome-like structures with equal length struts we leave open for exploration. Beyond the physics of crystalline defects, the experimental playground we have developed is ideal for exploring general questions concerning the physics of order and disorder<sup>25</sup> in curved space.

## METHODS SUMMARY

**Sample preparation.** The oil–glycerol surfaces coated with PMMA particles are prepared as follows. First, the glycerol droplets and capillary bridges are prepared in contact with air in a capillary channel. Then, the channel is filled with particles suspended in the oil. Particles in the bulk can be removed by subsequently flushing the channel with clean oil. The PMMA particles, prepared following refs 26, 27, are coated with a layer of poly(hydroxy stearic acid) (PHSA), which charges positively ( $\sim 100$  charges per particle) in the oil<sup>28</sup>. The oil phase is a mixture of cyclohexyl bromide (CHB) and dodecane that matches the refractive index of glycerol, allowing us to image with minimal distortion the full surface even when it is highly curved (Fig. 1). The particles are highly hydrophobic, but nonetheless bind to the glycerol–oil interface by image charge attraction to the higher dielectric glycerol ( $\epsilon_{\text{glycerol}} \approx 42$ ,  $\epsilon_{\text{CHB}} \approx 7.9$ )<sup>28</sup>. The dielectric contrast also drives the migration of ions from the oil phase into the glycerol, cleaning the oil above the interface, leading to a Debye screening length of  $\sim 50 \mu\text{m}$ . Furthermore, positive ions are pumped from the oil to the water phase preferentially over negative ions, causing the glycerol to charge positively and repel residual particles suspended in the oil. These effects combine to create a clean lattice of strongly repulsive particles uninfluenced by particles in the bulk.

**Imaging and reconstruction.** Particles were imaged using a Yokogawa CSU-10 spinning disk confocal microscope and a Leica SP5 confocal microscope. Confocal images were rendered using Andor IQ. Particles were tracked using standard IDL tracking routines<sup>29</sup>. Triangulations and rendering were done using custom codes written in Matlab.

Received 29 July; accepted 25 October 2010.

- Fuller, R. Building construction. US Patent US2682235, 1–12 (1954).
- Bausch, A. R. *et al.* Grain boundary scars and spherical crystallography. *Science* **299**, 1716–1718 (2003).
- Lipowsky, P., Bowick, M., Meinke, J., Nelson, D. & Bausch, A. Direct visualization of dislocation dynamics in grain boundary scars. Preprint at (<http://arxiv.org/abs/cond-mat/0506366>) (2005).
- Einert, T., Lipowsky, P., Schilling, J., Bowick, M. & Bausch, A. Grain boundary scars on spherical crystals. *Langmuir* **21**, 12076–12079 (2005).
- Nelson, D. & Peliti, L. Fluctuations in membranes with crystalline and hexatic order. *J. Phys.* **48**, 1085–1092 (1987).

- Pérez-Garrido, A., Dodgson, M. & Moore, M. Influence of dislocations in Thomson's problem. *Phys. Rev. B* **56**, 3640–3643 (1997).
- Bowick, M., Nelson, D. & Travesset, A. Interacting topological defects on frozen topographies. *Phys. Rev. B* **62**, 8738–8751 (2000).
- Vitelli, V., Lucks, J. B. & Nelson, D. R. Crystallography on curved surfaces. *Proc. Natl Acad. Sci. USA* **103**, 12323–12328 (2006).
- Santangelo, C., Vitelli, V., Kamien, R. & Nelson, D. Geometric theory of columnar phases on curved substrates. *Phys. Rev. Lett.* **99**, 017801 (2007).
- Giomì, L. & Bowick, M. Defective ground states of toroidal crystals. *Phys. Rev. E* **78**, 010601 (2008).
- Sausset, F., Tarjus, G. & Nelson, D. R. Structure and dynamics of topological defects in a glassy liquid on a negatively curved manifold. *Phys. Rev. E* **81**, 031504 (2010).
- Hexemer, A. *Order and Disorder of Block Copolymers and Particles on Surfaces with Topology*. PhD thesis, Univ. California Santa Barbara (2006).
- DeVries, G. A. *et al.* Divalent metal nanoparticles. *Science* **315**, 358–361 (2007).
- Chaikin, P. & Lubensky, T. *Principles of Condensed Matter Physics* (Cambridge Univ. Press, 1995).
- Nelson, D. *Defects and Geometry in Condensed Matter Physics* (Cambridge Univ. Press, 2002).
- Lipowsky, P., Bowick, M., Meinke, J., Nelson, D. & Bausch, A. Direct visualization of dislocation dynamics in grain-boundary scars. *Nature Mater.* **4**, 407–411 (2005).
- Bowick, M. J., Giomì, L., Shin, H. & Thomas, C. K. Bubble-raft model for a paraboloidal crystal. *Phys. Rev. E* **77**, 021602 (2008).
- Hexemer, A., Vitelli, V., Kramer, E. & Fredrickson, G. Monte Carlo study of crystalline order and defects on weakly curved surfaces. *Phys. Rev. E* **76**, 051604 (2007).
- Bowick, M. J. & Giomì, L. Two-dimensional matter: order, curvature and defects. *Adv. Phys.* **58**, 449–563 (2009).
- Nelson, D., Turner, A. M., Vitelli, V. & Nelson, D. R. Vortices on curved surfaces. *Rev. Mod. Phys.* **82**, 1301–1348 (2010).
- Kamien, R. The geometry of soft materials: a primer. *Rev. Mod. Phys.* **74**, 953–971 (2002).
- Gillette, R. & Dyson, D. Stability of fluid interfaces of revolution between equal solid circular plates. *Chem. Eng. J.* **2**, 44–54 (1971).
- Langbein, D. W. *Capillary Surfaces: Shape-Stability-Dynamics, in Particular Under Weightlessness* (Springer, 2002).
- Weertman, J. & Weertman, J. R. *Elementary Dislocation Theory* (Oxford Univ. Press, 1992).
- Modes, C. D. & Kamien, R. D. Hard disks on the hyperbolic plane. *Phys. Rev. Lett.* **99**, 235701 (2007).
- Bosma, G. *et al.* Preparation of monodisperse, fluorescent PMMA-latex colloids by dispersion polymerization. *J. Colloid Interface Sci.* **245**, 292–300 (2002).
- Antl, L. *et al.* The preparation of poly(methyl methacrylate) lattices in nonaqueous media. *Colloids Surf.* **17**, 67–78 (1986).
- Leunissen, M. E., van Blaaderen, A., Hollingsworth, A. D., Sullivan, M. T. & Chaikin, P. M. Electrostatics at the oil–water interface, stability, and order in emulsions and colloids. *Proc. Natl Acad. Sci. USA* **104**, 2585–2590 (2007).
- Crocker, J. & Grier, D. Methods of digital video microscopy for colloidal studies. *J. Colloid Interface Sci.* **179**, 298–310 (1996).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We acknowledge discussions with M. Bowick, A. Grosberg, S. Sacanna and A. M. Turner. W.T.M.I. acknowledges guidance in particle synthesis from A. D. Hollingsworth and M. T. Elsesser. W.T.M.I. acknowledges support from Rhodia and the English Speaking Union. P.M.C. acknowledges support from MRSEC DMR-0820341 and NASA NNX08AK04G. W.T.M.I. and V.V. acknowledge hospitality from Stichting FOM and the Aspen Center for Physics.

**Author Contributions** P.M.C. and W.T.M.I. initiated and designed research. W.T.M.I. designed and performed experiments, analysed data and synthesized colloidal particles. W.T.M.I., P.M.C. and V.V. interpreted data. V.V. and W.T.M.I. performed elasticity calculations. W.T.M.I. and P.M.C. wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to W.T.M.I. ([wtmirvine@uchicago.edu](mailto:wtmirvine@uchicago.edu)) and P.M.C. ([chaikin@nyu.edu](mailto:chaikin@nyu.edu)).

# Tidal dissipation and the strength of the Earth's internal magnetic field

Bruce A. Buffett<sup>1</sup>

Magnetic fields at the Earth's surface represent only a fraction of the field inside the core<sup>1</sup>. The strength and structure of the internal field are poorly known<sup>2–5</sup>, yet the details are important for our understanding of the geodynamo. Here I obtain an indirect estimate for the field strength from measurements of tidal dissipation. Tidally driven flow in the Earth's liquid core develops internal shear layers, which distort the internal magnetic field and generate electric currents. Ohmic losses damp the tidal motions and produce detectable signatures in the Earth's nutations. Previously reported evidence of anomalous dissipation in nutations<sup>3,6</sup> can be explained with a core-averaged field of 2.5 mT, eliminating the need for high fluid viscosity<sup>6</sup> or a stronger magnetic field at the inner-core boundary<sup>3</sup>. Estimates for the internal field constrain the power required for the geodynamo<sup>7,8</sup>.

Tidal forces with periods of around one cycle per day cause variations in the direction of the Earth's rotation (known as precession and nutation). The motion includes separate angular velocities for the solid and fluid cores ( $\Omega_s$  and  $\Omega_f$ , respectively), which are misaligned with the angular velocity of the mantle,  $\Omega_m$ . Coupling between the motion of the core and the mantle gives rise to free modes of oscillation that are resonantly excited at nearby tidal periods. One of these modes, the free core nutation, is characterized by angular velocities  $\Omega_s$  and  $\Omega_f$  that tilt together from the initial angular velocity of the mantle. The other mode is usually called the free inner-core nutation (FICN) because the motion mainly involves a tilt of  $\Omega_s$ .

Flow in the fluid core is often approximated by a solution due to Poincaré<sup>9</sup>. Fluid vorticity in the core is initially  $2\Omega_0\hat{z}$ , where  $\Omega_0$  is the rotation rate of the Earth and  $\hat{z}$  is the polar axis of the mantle. When  $\Omega_f$  tilts out of alignment with  $\hat{z}$ , the fluid is assumed to have uniform vorticity of  $2\Omega_f$ , corresponding to rigid-body rotation about an inclined axis. However, a rigid-body rotation cannot satisfy boundary conditions on the fluid velocity because of the spheroidal shape of the core. Consequently, the solution for the fluid velocity has the form

$$\mathbf{v}_f(\mathbf{r}, t) = [\Omega_f(t) - \Omega_m(t)] \times \mathbf{r} + \mathbf{v}'(\mathbf{r}, t)$$

where  $\mathbf{r}$  is the position vector,  $t$  is time and  $\mathbf{v}'$  describes the deviation from rigid-body rotation. Poincaré's solution for  $\mathbf{v}'$  enforces boundary conditions at the core–mantle boundary without disturbing the condition of uniform vorticity. Thus, an exact solution is found for the flow of an inviscid fluid when the core is entirely liquid (that is, there is no inner core). The solution for  $\mathbf{v}'$  can be represented by a potential flow, which is readily adapted to satisfy boundary conditions when the inner core is present. However, the resulting fluid velocity does not obey the governing equations when the inner core tilts out of alignment with the mantle (Methods). Strain in the flow alters the vorticity (through vortex stretching), invalidating the assumption of uniform vorticity.

The breakdown of Poincaré's solution is due primarily to a tilt of the inner core. I investigate the breakdown by determining  $\mathbf{v}'$  when the angular velocities  $\Omega_f$  and  $\Omega_s$  are prescribed. It suffices to assume that the mantle rotates with constant angular velocity,  $\Omega_0\hat{z}$ , while  $\Omega_s(t)$  and  $\Omega_f(t)$  vary periodically with frequency  $\omega$ . The amplitude and phase of  $\Omega_s(t)$  and  $\Omega_f(t)$  are chosen to conserve angular momentum, so the motion corresponds closely to the FICN free mode.

Differential rotation between the fluid and solid cores causes a small radial motion at the inner-core boundary due to hydrostatic flattening. The total fluid velocity,  $\mathbf{v}_f$ , is governed by the linearized Navier–Stokes equation

$$\frac{\partial \mathbf{v}_f}{\partial t} + 2\hat{z} \times \mathbf{v}_f = -\nabla P + E\nabla^2 \mathbf{v}_f$$

which is written in non-dimensional form using the radius,  $R$ , of the core as a length scale and  $\Omega_0^{-1}$  as a timescale;  $P$  is the fluid pressure and the Ekman number,  $E = \nu/\Omega_0 R^2$ , characterizes the effect of fluid viscosity,  $\nu$ , on the dynamics. For typical parameters<sup>10,11</sup> ( $\nu = 10^{-6} \text{ m}^2 \text{ s}^{-1}$ ,  $R = 3.48 \times 10^6 \text{ m}$ ,  $\Omega_0 = 0.73 \times 10^{-4} \text{ s}^{-1}$ ), the Ekman number is roughly  $10^{-15}$ . Numerical calculations for realistic values of  $E$  are not feasible, but it is possible to decrease  $E$  until a persistent pattern of flow emerges. Scaling relations are used to extrapolate the numerical solution to the appropriate value of  $E$ .

Figure 1a shows the kinetic energy in  $\mathbf{v}'$  on a meridional cross-section through the outer core for  $E = 2 \times 10^{-7}$ . The kinetic energy is concentrated in narrow bands (shear layers) that form conical surfaces in three dimensions. The orientation of the bands is set by the propagation and reflection of inertial waves; closed propagation paths within the core establish the location of the shear layers<sup>12,13</sup>.

The imposed radial motion at the inner-core boundary (maximum at mid latitudes) is not evident in Fig. 1a because the flow associated with the internal shear layers is much larger. Indeed, the maximum flow occurs at the rotation axis, where the wave motion converges along conical surfaces<sup>14</sup>.

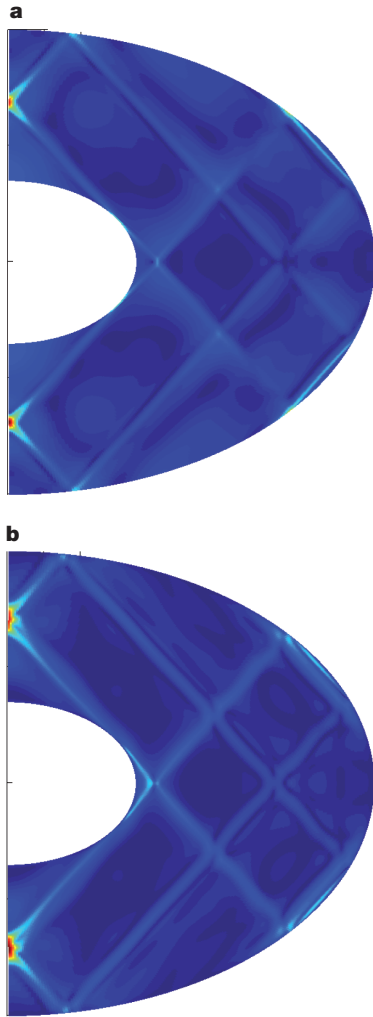
The width,  $l_v$ , of the shear layer decreases as  $E$  is lowered. A quantitative estimate for  $l_v$  is obtained from the ratio of viscous dissipation to kinetic energy in  $\mathbf{v}'$  (Supplementary Information). Once  $E < 10^{-6}$ , the width approaches an asymptotic value of  $l_v = E^{1/3}$ , which is consistent with theoretical estimates<sup>15,16</sup>. A realistic value of  $E = 10^{-15}$  yields  $l_v = 10^{-5}$ , corresponding to 35 m in the core.

Narrow shear layers distort the local magnetic field and generate electric currents. Most of the current is produced inside the shear layers, but diffusion allows the current to spread beyond them. The magnetic skin depth is nominally 200 m at diurnal periods, so the current layer is thicker than the shear layer. A quantitative description of magnetic perturbation,  $\mathbf{b}$ , is given by the linearized induction equation, which is written in non-dimensional form as

$$\frac{\partial \mathbf{b}}{\partial t} = \nabla \times (\mathbf{v}' \times \mathbf{B}) + D\nabla^2 \mathbf{b}$$

where  $\mathbf{B}$  is the initial magnetic field and  $D = \eta/\Omega_0 R^2$  characterizes the magnetic diffusivity,  $\eta$ . I set  $D = 10E^{2/3}$  so that the width of the current layer (or skin depth,  $l_b = (2D)^{1/2}$ ) is constant relative to  $l_v$  as  $E$  is varied, yielding the expected value when  $E = 10^{-15}$ . A uniform field,  $\mathbf{B} = B\hat{z}$ , is assumed for simplicity because the structure of the internal field is unknown. Larger perturbations occur if  $\mathbf{B}$  is perpendicular to the shear layer, but this configuration is unlikely to be typical. The intersection of shear layers with a uniform, vertical field approximates the average dissipation for a random field orientation, which may be appropriate

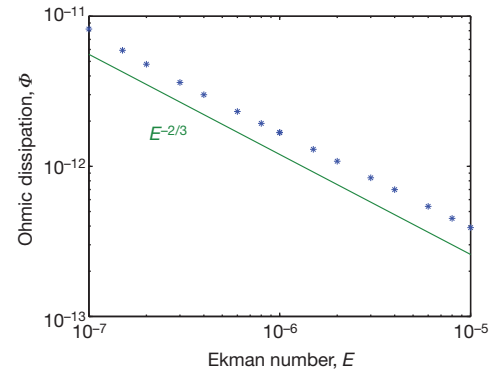
<sup>1</sup>Department of Earth and Planetary Science, University of California, Berkeley, California 94720, USA.



**Figure 1 | Structure of flow and magnetic field perturbation in the fluid core when the inner core tilts out of alignment with the mantle.** **a**, Kinetic energy in  $\mathbf{v}'$  on a meridional cross-section for  $E = 2 \times 10^{-7}$ . The flow is driven by radial motion at the inner-core boundary with a prescribed frequency  $\omega = \Omega_0(1 - \varepsilon)$ , where  $\varepsilon = 0.0025$  is the hydrostatic flattening of the inner core. Shear layers of width  $O(E^{1/3})$  are oriented in the direction of inertial wave propagation. Red, highest energy; deep blue, reference (zero-energy) state. **b**, Magnetic energy in the perturbed field due to the influence of shear layers on a uniform, vertical magnetic field. Steep gradients in the perturbed magnetic field cause electric currents that damp the tidal motion.

for the true field. Figure 1b shows the magnetic energy in **b** for  $E = 2 \times 10^{-7}$  and an initial magnetic field of  $B = 1$  mT. A comparison of the ohmic dissipation with the magnetic energy in **b** confirms that  $l_b$  is approximately equal to the skin depth.

The amplitude of **b**,  $b$ , depends on both the shear,  $\mathbf{v}'/l_v$ , and the ratio of length scales  $l_b/l_v$ . Because the ratio  $l_b/l_v$  is held constant in the calculation, the magnetic perturbation is expected to scale as  $b/B \propto \nu/E^{1/3}$ , where  $\nu$  is the amplitude of  $\mathbf{v}'$ . Order-of-magnitude estimates for the resulting magnetic energy and ohmic dissipation are  $b^2$  and  $Db^2/l_b^2$ , respectively. Both of these quantities vary with Ekman number as  $E^{-2/3}$  because  $l_b$  is set by the skin depth. Direct calculations confirm the trend in dissipation with  $E$  (Fig. 2). Thus, the ohmic dissipation is substantially larger for realistic values of  $E$  owing to the thin shear layers. For example, the dissipation rate for  $E = 10^{-7}$  in Fig. 2 increases to  $\Phi = 1.8 \times 10^{-6}$  at  $E = 10^{-15}$  for an internal magnetic field of 1 mT. By comparison, the (dimensionless) kinetic energy of the motion is  $4.4 \times 10^{-3}$ . This level of dissipation is sufficient to explain an anomalous source of dissipation in the Earth's nutations.



**Figure 2 | Ohmic dissipation in the core due to internal shear layers.**

Decreasing the Ekman number  $E$  causes higher ohmic dissipation: a series of calculations confirm the expected dependence,  $\Phi \propto E^{-2/3}$ , which is indicated by the slope of the solid line. The quality factor is calculated from equation (1) using  $2E_k = A_s |\mathbf{\Omega}_s - \mathbf{\Omega}_m|^2$ , where  $A_s$  is the moment of inertia of the inner core. The numerical calculations use a constant-amplitude tilt of  $|\mathbf{\Omega}_s - \mathbf{\Omega}_m| = \Omega_0$ , corresponding to a unit-amplitude tilt in non-dimensional variables. The non-dimensional kinetic energy is  $E_k = 4.4 \times 10^{-3}$ .

Dissipation is evident in nutation measurements as a phase lag between the tidal force and the Earth's response. Corrections are routinely applied for the effects of mantle anelasticity<sup>17</sup> and ocean tides<sup>18</sup>, but several large phase lags remain unexplained, particularly at tidal periods close to the natural periods of the free modes. The measured phase lags are often related to the quality factor,  $Q$ , of the free modes, in analogy to the relationship for a damped harmonic oscillator. A low  $Q$  for the FICN mode is inferred by fitting nutation measurements to a theoretical model<sup>19</sup>. An estimate<sup>3</sup> based on nutation measurements before 2000 yielded  $Q = 677$ , whereas a more recent study<sup>6</sup> using an extra nine years of data and a different estimation procedure obtained  $Q = 459 \pm 27$ . The quality factor for the predicted flow is given by

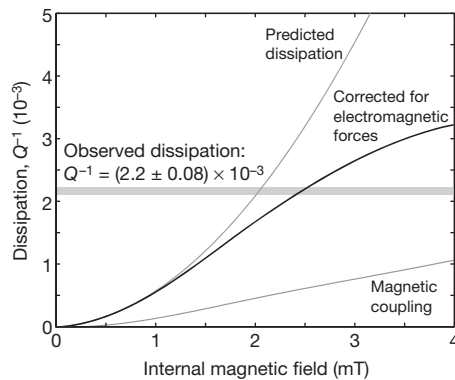
$$Q^{-1} = \frac{\tau \Phi}{2\pi E_k} \quad (1)$$

where  $\tau = 2\pi/\text{Re}(\omega)$  is the period of the motion and  $E_k$  is the total kinetic energy. Most of the kinetic energy in the FICN mode is due to relative rotation of the inner core (that is,  $\mathbf{\Omega}_s - \mathbf{\Omega}_m$ ). The dependences of  $E_k$  and  $\Phi$  on  $\mathbf{\Omega}_s - \mathbf{\Omega}_m$  are identical, so the predicted quality factor is independent of the amplitude of the inner-core tilt.

Figure 3 shows the predicted  $Q^{-1}$  as a function of the internal field strength,  $B$ . The numerical calculations represent a 'weak-field' approximation because the electromagnetic force due to **b** is not included in the dynamics. This approximation is reasonable when the field is weak, but it tends to overestimate the dissipation when the field is strong<sup>20</sup>. Scaling arguments suggest that the electromagnetic force is small in comparison with the viscous force at computationally accessible values of  $E$ , so a reasonable internal field would have little influence on the calculated flow. Instead of including the electromagnetic force in the calculation, I correct the numerical results using a local analysis for the interaction of flow and magnetic field in the shear layers.

The shear layers can be represented by a linear superposition of inertial waves with a short wavelength in the direction perpendicular to the layers. The amplitude of the local wavenumber is nominally  $k \approx l_v^{-1}$ , which characterizes the width of the layers. A plane-wave solution for the inertial waves can be coupled to the magnetic induction equation to account explicitly for the electromagnetic force in the dynamics of the waves (Supplementary Information). The local plane-wave solution permits calculations with realistic values for viscosity and field strength, avoiding some of the limitations of the numerical calculations. Including the electromagnetic force in the dynamics decreases the amplitude of the waves, reducing both the electric current and the associated dissipation. These local solutions are used to obtain estimates for the ohmic dissipation with and without the





**Figure 3 | Predicted dissipation as a function of magnetic field strength.**

The numerical calculation represents a weak-field approximation because the electromagnetic force is not included in the dynamics. A correction for the electromagnetic force is based on a local analysis of inertial waves in the presence of a magnetic field. Ohmic dissipation due to magnetic coupling at the inner-core boundary requires a radial magnetic field of 7 mT or more to explain the observed dissipation<sup>2</sup>. Internal shear layers in the outer core explain the observed dissipation with an average magnetic field of 2.5 mT.

addition of the electromagnetic force. The ratio of dissipation estimates is used to correct the numerical results in Fig. 3.

I also calculate the dissipation due to magnetic coupling at the inner-core boundary<sup>20</sup>. In this case, the strength of the magnetic field refers to the radial component at the inner-core boundary. A strong radial field is required to explain the observed dissipation with magnetic coupling<sup>3</sup>. Alternatively, a high fluid viscosity ( $\nu = 10 \text{ m}^2 \text{ s}^{-1}$ ) has been proposed<sup>6</sup> as the source of dissipation. Allowing for the influence of internal shear layers eliminates the need for a high viscosity or a strong radial field at the inner-core boundary. The nutation observations can be explained with a 2.5-mT field when the flow is corrected for the effects of electromagnetic forces. The weak-field limit establishes a lower bound of roughly 2 mT.

Numerical models<sup>1,21</sup> and theoretical consideration<sup>22</sup> suggest an internal magnetic field of 1–4 mT. The field strength inferred from tidal dissipation is compatible with these predictions, although the tidal estimate is not sensitive to the azimuthal component of the magnetic field. This lack of sensitivity is due to the conical structure of the shear layers, which causes little distortion of the azimuthal field. However, the tidal estimate is sensitive to the field that controls the propagation of torsional oscillations<sup>3</sup>. Recent evidence<sup>5</sup> for a strong field from the propagation of torsional oscillations is consistent with the estimate presented here. A source of uncertainty in the tidal estimate arises from the viscosity of the fluid core because it controls the thickness of the shear layers. A higher viscosity reduces the shear and decreases the ohmic dissipation. However, the change in the ratio  $I_b/I_v$  causes an increase in dissipation. The net effect is a modest decrease in dissipation. Doubling the viscosity requires a 10% increase in the field strength to explain the observed dissipation. In spite of these uncertainties, there are few observations that constrain the strength of the internal field. The tidal estimate represents a core-wide average with sensitivity to all but the azimuthal component of the field. It is striking that radio emissions from distant quasars<sup>23</sup> offer insights into the internal magnetic field by providing a precise determination of the Earth's nutations.

## METHODS SUMMARY

The numerical calculations are based on the dynamo model of ref. 24 with the nonlinear terms omitted. The velocity and magnetic fields are represented using vector spherical harmonics in a series expansion up to degree 260. I apply stress-free boundary conditions to the velocity field and insulating boundary conditions to the magnetic field. This choice ensures that all of the viscous and ohmic dissipation

is due to the internal shear layers. Use of no-slip boundary conditions introduces thinner,  $O(E^{1/2})$ , shear layers at the boundaries and a radial flow due to Ekman pumping, but the net effect on the dissipation is small for realistic values of  $E$ . I use finite differences for the radial derivatives on a uniformly spaced grid with up to 360 radial levels. Both the velocity and magnetic fields have periodic time dependence with a prescribed frequency. The governing equations yield a system of algebraic equations for the radial coefficients of the vector spherical harmonic expansion. I obtained solutions iteratively using the GMRES method<sup>25</sup>.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 28 July; accepted 1 November 2010.**

1. Aubert, J., Labrosse, S. & Poitou, C. Modelling the palaeo-evolution of the geodynamo. *Geophys. J. Int.* **179**, 1414–1428 (2009).
2. Zatman, S. & Bloxham, J. Torsional oscillations and the magnetic field within the Earth's core. *Nature* **388**, 760–763 (1997).
3. Mathews, P. M., Herring, T. A. & Buffett, B. A. Modeling of nutation and precession: new nutation series for nonrigid Earth and insights into the Earth's interior. *J. Geophys. Res.* **107**, 2068 (2002).
4. Buffett, B. A., Mound, J. & Jackson, A. Inversion of torsional oscillations for the structure and dynamics of Earth's core. *Geophys. J. Int.* **177**, 878–890 (2009).
5. Gillet, N., Jault, D., Canet, E. & Fournier, A. Fast torsional waves and strong magnetic field within the Earth's core. *Nature* **465**, 74–77 (2010).
6. Koot, L., Dumberry, M., Rivoldini, A., de Viron, O. & Dehant, V. Constraints on the coupling at the core-mantle and inner-core boundaries inferred from nutation observations. *Geophys. J. Int.* **182**, 1279–1294 (2010).
7. Jackson, A. & Livermore, P. On ohmic heating in the Earth's core: nutation constraints. *Geophys. J. Int.* **177**, 367–382 (2009).
8. Christensen, U. R. & Tilgner, A. Power requirements of the geodynamo from ohmic losses in numerical and laboratory dynamos. *Nature* **429**, 169–171 (2004).
9. Poincaré, H. Sur la précession des corps déformables. *Bull. Astron.* **27**, 321–356 (1910).
10. de Wijs, G. A. *et al.* The viscosity of liquid iron at the physical conditions of the Earth's core. *Nature* **392**, 805–807 (1998).
11. Dobson, D. P. Self-diffusion in liquid Fe at high pressure. *Phys. Earth Planet. Inter.* **130**, 271–284 (2002).
12. Hollerbach, R. & Kerswell, R. R. Oscillatory internal shear layers in rotating and precessing flows. *J. Fluid Mech.* **298**, 327–339 (1995).
13. Tilgner, A. Driven inertial oscillations in spherical shells. *Phys. Rev. E* **59**, 1789–1794 (1999).
14. Rieutord, M. & Valdettaro, L. Inertial waves in a rotating spherical shell. *J. Fluid Mech.* **341**, 77–99 (1997).
15. Walton, I. C. Viscous shear layers in an oscillating rotating fluid. *Proc. R. Soc. Lond. A* **344**, 101–110 (1975).
16. Kerswell, R. R. On the internal shear layers spawned by the critical regions in oscillatory Ekman boundary layers. *J. Fluid Mech.* **298**, 311–325 (1995).
17. Wahr, J. M. & Bergen, Z. The effects of mantle anelasticity on nutations, Earth tides and tidal variations in the rotation rate. *Geophys. J. R. Astron. Soc.* **87**, 633–668 (1986).
18. Desai, S. & Wahr, J. M. Empirical ocean tide models estimated TOPEX/Poseidon altimetry. *J. Geophys. Res.* **100**, 25205–25228 (1995).
19. Mathews, P. M., Buffett, B. A., Herring, T. A. & Shapiro, I. I. Forced nutations of the Earth: influence of inner core dynamics. 1. Theory. *J. Geophys. Res.* **96**, 8219–8242 (1991).
20. Buffett, B. A., Mathews, P. M. & Herring, T. A. Modeling of nutation and precession: effects of electromagnetic coupling. *J. Geophys. Res.* **107**, 2070 (2002).
21. Christensen, U. R. & Aubert, J. Scaling properties of convection-driven dynamos in rotating spherical shells and application to planetary magnetic fields. *Geophys. J. Int.* **166**, 97–114 (2006).
22. Starchenko, S. V. & Jones, C. A. Typical velocities and magnetic field strengths in planetary interiors. *Icarus* **157**, 426–435 (2002).
23. Sovers, O. J., Fenselow, J. L. & Jacobs, C. S. Astrometry and geodesy with radio interferometry: experiments, models, results. *Rev. Mod. Phys.* **70**, 1393–1454 (1998).
24. Kuang, W. & Bloxham, J. Numerical modeling of magnetohydrodynamic convection in a rapidly rotating spherical shell: weak and strong field dynamo action. *J. Comput. Phys.* **153**, 51–81 (1999).
25. Saad, Y. & Schultz, M. H. GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput.* **7**, 856–869 (1986).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The author declares no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to the author ([bubuffett@berkeley.edu](mailto:bubuffett@berkeley.edu)).

## METHODS

Poincaré's solution for the flow of an inviscid fluid can be represented by<sup>26</sup>

$$\mathbf{v}_f = \boldsymbol{\chi}_f \times \mathbf{r} + \nabla\psi \quad (2)$$

where  $\boldsymbol{\chi}_f = \boldsymbol{\Omega}_f - \boldsymbol{\Omega}_m$  is the relative angular velocity of the fluid and  $\psi$  is a scalar potential that describes the deviation from rigid-body rotation. The continuity condition,  $\nabla \cdot \mathbf{v}_f = 0$ , requires

$$\nabla^2\psi = 0$$

which is subject to the boundary condition

$$\mathbf{v}_f \cdot \mathbf{n} = 0 \quad (3)$$

on the normal component of the velocity at the core–mantle boundary. The spheroidal boundary is described in Cartesian coordinates  $x$ ,  $y$  and  $z$  by

$$\frac{x^2}{a^2} + \frac{y^2}{a^2} + \frac{z^2}{c^2} = 1$$

where  $a$  and  $c$  are the equatorial and polar radii, respectively. The flattening of the boundary is defined by

$$\varepsilon_f = \frac{a - c}{a}$$

and the outward normal,  $\mathbf{n}$ , can be approximated by

$$\mathbf{n} = \frac{x}{a}\hat{\mathbf{x}} + \frac{y}{a}\hat{\mathbf{y}} + \frac{z}{a}(1 + 2\varepsilon_f)\hat{\mathbf{z}}$$

to first order in  $\varepsilon_f$ .

The relative angular velocity is confined to the equatorial plane for small-amplitude motion. Consequently, I express  $\boldsymbol{\chi}_f$  in the form

$$\boldsymbol{\chi}_f = \tilde{\chi}_f(\hat{\mathbf{x}} + i\hat{\mathbf{y}})e^{i\omega t} \quad (4)$$

where  $\omega$  is the frequency of the motion. The sign convention in equation (4) requires  $\omega \approx +\Omega_0$  for diurnal tidal motions. Substituting equation (2) into the boundary condition in equation (3) yields

$$\nabla\psi \cdot \mathbf{n} = -2\varepsilon_f \tilde{\chi}_f(yz - ixz)a^{-1}e^{i\omega t} \quad (5)$$

for the normal component of the potential flow. The potential that satisfies equation (5) is

$$\psi = -\varepsilon_f \tilde{\chi}_f(yz - ixz)e^{i\omega t} \quad (6)$$

where second-order terms in  $\varepsilon_f$  are neglected.

This solution also satisfies boundary conditions at the inner-core boundary when the relative angular velocity of the inner core,  $\boldsymbol{\chi}_s = \boldsymbol{\Omega}_s - \boldsymbol{\Omega}_m$ , equals zero and the flattening,  $\varepsilon_s$ , of the inner-core boundary equals  $\varepsilon_f$ . (The difference in flattening is about 5% for a hydrostatic Earth<sup>27</sup>.) Departure from either of these conditions requires a correction to the potential flow. In principle, the potential flow can be adapted to satisfy boundary conditions at both the core–mantle boundary and the inner-core boundary. However, the resulting flow no longer satisfies the governing equations for an inviscid, rotating fluid.

The difficulty arises from the structure of the potential flow. The form of the solution in equation (2) ensures that the fluid has uniform vorticity because the potential flow has no vorticity. The solution for  $\psi$  in equation (6) is linear in  $z$ , which means that the vertical velocity is constant and the vertical strain rate,  $\partial^2\psi/\partial z^2$ , is zero. Thus the potential flow does not alter the uniform vorticity by stretching the dominant  $z$  component associated with steady rotation. In other

words, the form of the potential flow is compatible with the assumption of uniform vorticity in the flow. However, when the potential flow is modified to accommodate more general boundary conditions at the surface of the inner core, the resulting potential flow is no longer linear in  $z$ . Vertical strain due to the potential flow stretches the  $z$  component of vorticity, causing departures from uniform vorticity. As a result, the potential flow is now incompatible with the assumption of uniform vorticity. This means that a potential flow is insufficient to describe the deviations from rigid-body rotation in the more general case when a solid inner core is present.

A more general solution for flow in the core is represented by

$$\mathbf{v}_f = \boldsymbol{\chi}_f \times \mathbf{r} + \mathbf{v}'$$

where the deviation,  $\mathbf{v}'$ , from rigid-body rotation is expanded in vector spherical harmonics<sup>24</sup>. The boundary condition on the normal component of velocity at the inner-core boundary requires that

$$\mathbf{v}' \cdot \mathbf{n} = 2[\tilde{\chi}_s \varepsilon_s + \tilde{\chi}_f(\varepsilon_f - \varepsilon_s)](yz - ixz)a_s^{-1} \quad (7)$$

where  $a_s$  is the equatorial radius of the inner core and  $\tilde{\chi}_s$  is the amplitude of  $\boldsymbol{\chi}_s$ . The difference  $\varepsilon_f - \varepsilon_s$  is small in comparison with  $\varepsilon_s$ , and the amplitude of  $\boldsymbol{\chi}_f$  is small in comparison with that of  $\boldsymbol{\chi}_s$  when the angular momentum of the fluid and solid core is conserved. To focus on the role of the inner core in the numerical calculations, I set  $\varepsilon_f = 0$  and approximate equation (7) by

$$\mathbf{v}' \cdot \hat{\mathbf{r}} = 2[\tilde{\chi}_s - \tilde{\chi}_f]\varepsilon_s(yz - ixz)a_s^{-1}$$

where  $\mathbf{n} = \hat{\mathbf{r}}$  is valid to leading order in  $\varepsilon_s$ . In addition, I use stress-free boundary conditions at both the inner-core boundary and the core–mantle boundary and set  $\mathbf{v}' \cdot \hat{\mathbf{r}} = 0$  at the core–mantle boundary. The non-zero radial motion at the inner-core boundary requires care in the definition of stress-free boundary conditions.

Both  $\boldsymbol{\chi}_f$  and  $\boldsymbol{\chi}_s$  are prescribed in the numerical calculations for  $\mathbf{v}'$ . The amplitude and phase are chosen to conserve angular momentum in the core, and the frequency of the motion is  $\omega = \Omega_0(1 - \varepsilon_s)$ , corresponding to the frequency of the FICN in the absence of elastic deformation and electromagnetic coupling at the fluid–solid boundaries.

The ratio of dissipation to kinetic energy (or  $Q^{-1}$ ) is independent of the amplitude of the imposed rotation, so I set  $\tilde{\chi}_s = 1$  and  $\tilde{\chi}_f = -A_s/A_f$  where  $A_s$  and  $A_f$  are the moments of inertia of the solid and fluid cores, respectively (assuming constant density in the fluid and solid). Similar results for the shear flow and dissipation are obtained if a value for  $\tilde{\chi}_s$  is imposed and the total velocity,  $\mathbf{v}_f$  (rather than  $\mathbf{v}'$ ), is expanded in vector spherical harmonics. In this case, the numerical solution includes the rigid rotation of the fluid core. Conservation of angular momentum can be used as a check on the calculations at low values of  $E$ . However, the flow associated with the average rotation of the fluid is larger by a factor of roughly  $1/\varepsilon_s$  than the deviation about a rigid rotation. Consequently, it is preferable to solve directly for the deviation,  $\mathbf{v}'$ , using imposed values for  $\boldsymbol{\chi}_s$  and  $\boldsymbol{\chi}_f$ . In either case, a rigid rotation is removed from the flow before solving the induction equation for the magnetic perturbation because the main magnetic field is specified in a frame that rotates with the fluid. The use of electrically insulating boundary conditions in the solution of the induction equation ensures that magnetic perturbations are due solely to the effect of shear layers in the flow.

26. Tilgner, A. in *Treatise on Geophysics* Vol. 8 (ed. Schubert, G.) 207–243 (Elsevier, 2007).

27. Mathew, P. M., Buffett, B. A., Herring, T. A. & Shapiro, I. I. Forced nutations of the Earth: influence of inner core dynamics, 2. Results. *J. Geophys. Res.* **96**, 8219–8242 (1991).

# Greenhouse gas mitigation can reduce sea-ice loss and increase polar bear persistence

Steven C. Amstrup<sup>1†</sup>, Eric T. DeWeaver<sup>2</sup>, David C. Douglas<sup>3</sup>, Bruce G. Marcot<sup>4</sup>, George M. Durner<sup>1</sup>, Cecilia M. Bitz<sup>5</sup> & David A. Bailey<sup>6</sup>

On the basis of projected losses of their essential sea-ice habitats, a United States Geological Survey research team concluded in 2007 that two-thirds of the world's polar bears (*Ursus maritimus*) could disappear by mid-century if business-as-usual greenhouse gas emissions continue<sup>1–3</sup>. That projection, however, did not consider the possible benefits of greenhouse gas mitigation. A key question is whether temperature increases lead to proportional losses of sea-ice habitat, or whether sea-ice cover crosses a tipping point and irreversibly collapses when temperature reaches a critical threshold<sup>4–6</sup>. Such a tipping point would mean future greenhouse gas mitigation would confer no conservation benefits to polar bears. Here we show, using a general circulation model<sup>7</sup>, that substantially more sea-ice habitat would be retained if greenhouse gas rise is mitigated. We also show, with Bayesian network model outcomes, that increased habitat retention under greenhouse gas mitigation means that polar bears could persist throughout the century in greater numbers and more areas than in the business-as-usual case<sup>3</sup>. Our general circulation model outcomes did not reveal thresholds leading to irreversible loss of ice<sup>6</sup>; instead, a linear relationship between global mean surface air temperature and sea-ice habitat substantiated the hypothesis that sea-ice thermodynamics can overcome albedo feedbacks proposed to cause sea-ice tipping points<sup>5,6,8</sup>. Our outcomes indicate that rapid summer ice losses in models<sup>9</sup> and observations<sup>6,10</sup> represent increased volatility of a thinning sea-ice cover, rather than tipping-point behaviour. Mitigation-driven Bayesian network outcomes show that previously predicted declines in polar bear distribution and numbers<sup>3</sup> are not unavoidable. Because polar bears are sentinels of the Arctic marine ecosystem<sup>11</sup> and trends in their sea-ice habitats foreshadow future global changes, mitigating greenhouse gas emissions to improve polar bear status would have conservation benefits throughout and beyond the Arctic<sup>12</sup>.

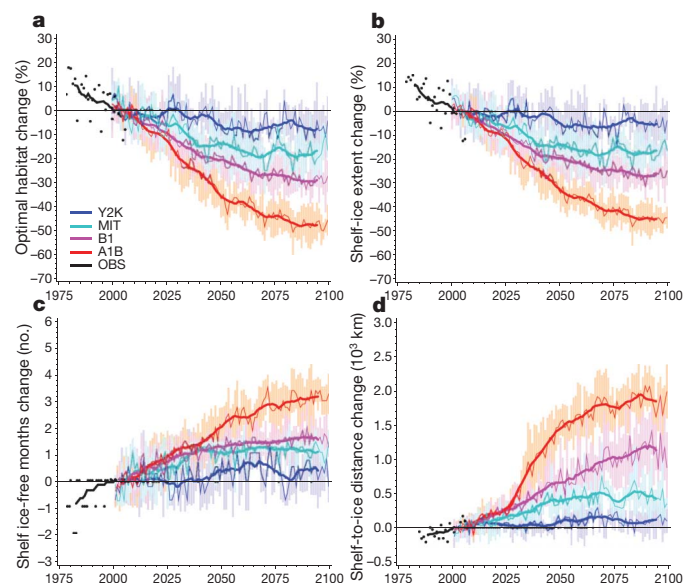
Polar bears are dependent on the sea ice for access to their marine mammal prey<sup>13,14</sup>, and occur only in Northern Hemisphere marine areas that are ice covered for long enough periods to allow sufficient foraging opportunity. Observed declines in summer sea ice have been associated with declining physical stature and condition, poorer survival and declining population size<sup>2,15,16</sup>. The anticipated future loss of sea-ice habitats resulting from global warming<sup>1</sup> was the principal driver of polar bear declines projected by the United States Geological Survey (USGS) studies<sup>3,17</sup>. Improved management of hunting and other human activities was found unable to materially alter this outcome (see plate 6 in ref. 3).

The USGS studies relied on general circulation model (GCM)-projected losses of Arctic sea ice based on the Special Report on Emissions Scenarios (SRES)<sup>18</sup> A1B 'business as usual' greenhouse gas emissions scenario. Recent emissions trends make it clear that without mitigation little departure from the 2007 polar bear projections could be expected<sup>19</sup>. Also, the hypothesis that the climate system contains tipping

elements<sup>4</sup> means that habitats supporting cold-dependent species could disappear abruptly and irreversibly when a particular global mean surface air temperature (GMAT) is exceeded<sup>6</sup>. It has been proposed<sup>20</sup> that existing greenhouse gas emissions already have committed the earth to temperatures that will rise above the tipping point for loss of perennial Arctic sea ice. The perception that nothing can be done to avoid catastrophic losses and ultimate disappearance of polar bears was exemplified in 2007 when the general media proclaimed polar bears were irreversibly doomed<sup>21</sup>.

We used projections of twenty-first century GMAT and sea-ice extent from the Community Climate System Model version 3 (CCSM3)<sup>7</sup> to test the hypothesis that a tipping point<sup>4–6,8</sup> will lead to irreversible loss of sea-ice habitats as GMAT increases. We used a Bayesian network model<sup>3</sup> to evaluate whether mitigating greenhouse gas rise could improve the future outlook for polar bears compared to previous projections.

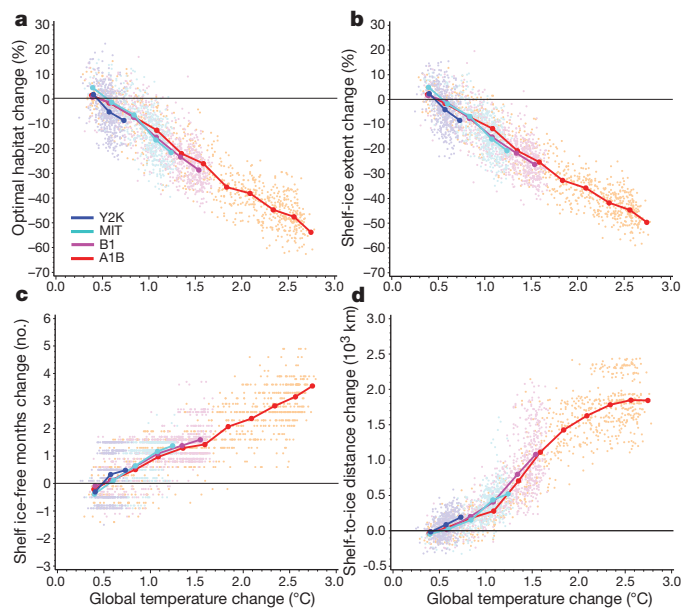
CCSM3 simulations were forced with greenhouse gas concentrations from five emissions scenarios (Supplementary Table 1): SRES<sup>18</sup> A1B and B1; the 2000 (Y2K) climate change commitment scenario<sup>22</sup>;



**Figure 1 | Changes from the present in polar bear habitat features varied greatly among greenhouse gas scenarios.** a–d, The DIV is illustrated here. Shown are changes in optimal polar bear foraging habitat (a), extent of sea ice over continental shelves (b), number of months continental shelves are ice free (c) and the distance from the shelf edge to the edge of the perennial pack ice as projected by CCSM3 with four greenhouse gas scenarios (defined in text) (d). Thin lines plot annual averages of the model runs under each greenhouse gas scenario, with error bars showing data  $\pm 1$  s.d. Bold lines are 10-year centred running averages of the annual mean values. OBS is observed passive microwave satellite data, black dots are the annual satellite observed values.

<sup>1</sup>US Geological Survey, Alaska Science Center, 4210 University Drive, Anchorage, Alaska 99508, USA. <sup>2</sup>National Science Foundation, 4201 Wilson Blvd., Arlington, Virginia 22230, USA. <sup>3</sup>US Geological Survey, Alaska Science Center, 3100 National Park Road, Juneau, Alaska 99801, USA. <sup>4</sup>USDA Forest Service, PNW Research Station, 620 SW Main St., Suite 400, Portland, Oregon 97205, USA. <sup>5</sup>Atmospheric Sciences, University of Washington, Seattle, Washington 98195, USA. <sup>6</sup>National Center for Atmospheric Research, 1850 Table Mesa Dr., Boulder, Colorado 80305, USA. <sup>†</sup>Present address: Polar Bears International, 810 N. Wallace, Suite E, P. O. Box 3008, Bozeman, Montana 59722, USA.

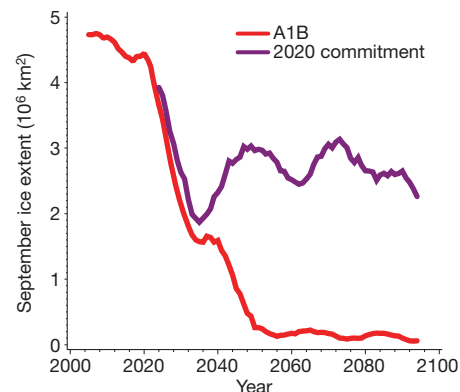




**Figure 2 | Relationship between GMAT change and change in polar bear habitat features is essentially linear.** a–d, The DIV is illustrated here. The optimal polar bear foraging habitat (a), extent of sea ice over continental shelves (b), number of months continental shelves are ice free (c) and the distance from the shelf edge to the edge of the perennial pack ice (d). Linear relationship between habitat and GMAT changes does not support the tipping-point hypothesis. Projections are from CCSM3 running four different greenhouse gas scenarios (defined in text).

the Level 1 stabilization scenario (CCSP450) of the United States Climate Change Science Program<sup>23</sup>; and the alternative scenario (AS)<sup>24</sup>. We pooled the AS and CCSP450 realizations into a 5-run mitigation (MIT) ensemble.

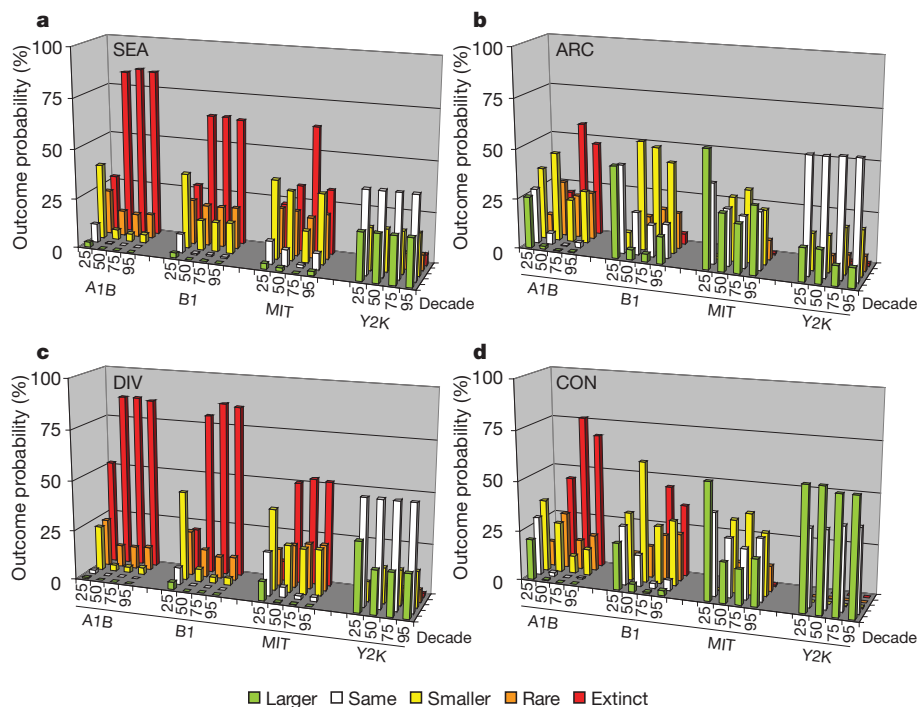
Reduced radiative forcing with greenhouse gas mitigation resulted in cooler temperatures, greater sea-ice retention (Supplementary Figs 3



**Figure 3 | September sea-ice extent (50% concentration) recovers from a RILE in a 2020 greenhouse gas commitment realization.** In the 2020 commitment realization, which was integrated from the same initial state as the A1B reference realization, greenhouse gas concentrations followed the A1B scenario until 2020, and were fixed thereafter. RILEs occurred in both realizations during the decade of the 2020s. In contrast to the reference run (red line), the substantial sea-ice recovery in the 2020 commitment scenario (purple line) supports the concept that RILEs represent natural sea-ice variability superimposed on a secular warming-induced sea-ice decline, rather than tipping points. All lines represent 10-year running averages compiled from the annual data.

and 4) and less change in important polar bear habitat features (Fig. 1). Importantly, the relationship between GMAT and projected habitat change was largely linear (Fig. 2). Even in September, the month of minimum ice cover, as GMAT increased sea ice and polar-bear-habitat availability smoothly decreased—regardless of the greenhouse gas scenario (Supplementary Figs 5 and 6).

We rejected the null hypothesis that there is a tipping point<sup>4–6,8</sup> of perennial Arctic sea-ice collapse by our failure to find a critical temperature threshold in our GCM outcomes. Our model outcomes support the alternative hypothesis that sea-ice thermodynamics can



**Figure 4 | Future polar bear persistence varies among ecoregions and greenhouse gas scenarios.** Bayesian network model projected outcomes (coloured bars) are shown for each of four greenhouse gas scenarios, four future decades, and four ecoregions. Although substantial risk of extirpation

continues for the SEA and DIV even with mitigation, increased levels of greenhouse gas mitigation improve the probability of future polar bear persistence in all ecoregions. In the x-axis legend, we refer to the decades of 2020–2029, 2045–2054, 2070–2079 and 2090–2099 as years 25, 50, 75 and 95.

dominate and reduce the destabilizing effects of the ice-albedo feedback on summer sea-ice cover<sup>6,25,26</sup>.

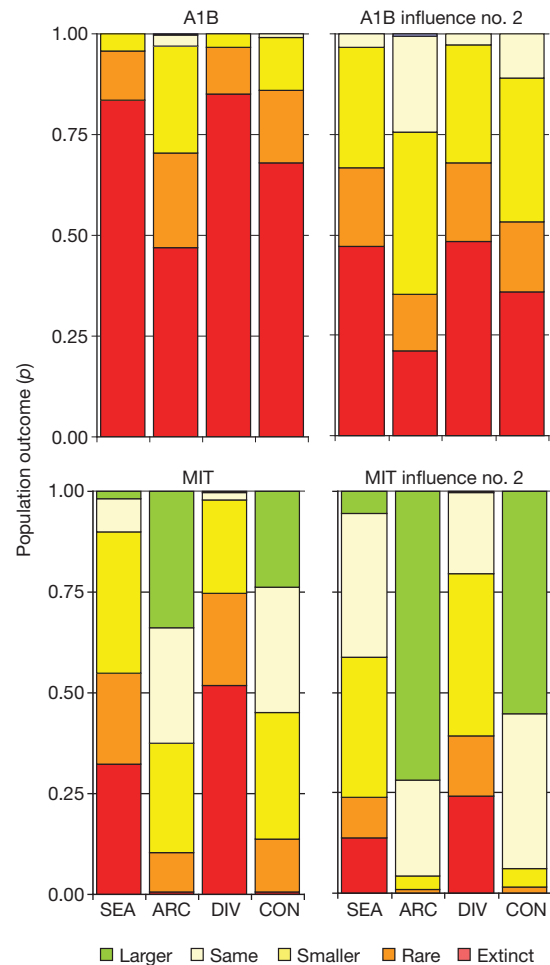
To test further for evidence of tipping-point behaviour, we compared rapid ice-loss events (RILEs)<sup>9,27</sup> in CCSM3 realizations using A1B<sup>18</sup> greenhouse gas levels and levels from a 2020 commitment integration in which greenhouse gas concentrations followed A1B until 2020 and were fixed at 2020 levels thereafter. In the A1B reference run, a RILE occurred between 2020 and 2030, and September Arctic sea ice largely disappeared by mid-century (Fig. 3). If RILEs represent tipping-point behaviour, as suggested<sup>5</sup>, the 2020 commitment run should have shown either no RILE or the same kind of permanent ice loss following a RILE as the reference run—depending on whether the climate system in that realization crossed the tipping point.

A RILE did occur in the 2020 commitment run. Instead of proceeding towards permanent ice loss as in the reference run, however, the RILE in the 2020 commitment run was followed by partial recovery and substantial retention of September sea-ice cover through the century (Fig. 3). Because the 2020 commitment run was integrated from the same 2020 initial state as the A1B reference, it experienced the same near-term natural variability, including a RILE during the 2020s. The 2020 commitment run did not proceed to an irreversible and unstoppable loss of remaining ice<sup>6</sup>, presumably because the long-term ice loss in CCSM3 is dictated by greenhouse gas radiative forcing and consequent global warming, which are substantially lower for the 2020 commitment run than A1B. This outcome indicates that RILEs are caused by the increased volatility of a thinner and more sensitive sea-ice cover, rather than the sea ice crossing an albedo-induced threshold from which it cannot return<sup>9,28,29</sup>.

The linear relationship between GMAT and sea-ice habitat change, and the return of sea ice after the RILE in our 2020 commitment experiment confirm that there is no tipping point<sup>4–6,8</sup> for summer Arctic sea ice in the CCSM3 climate model. We recognize that the absence of tipping points in a climate model does not guarantee that tipping-point behaviour will not occur in the real world. We recognize also that absence of tipping-point behaviour in one GCM does not necessarily mean that tipping points would not be present in other GCMs. Because sea-ice loss in CCSM3 is more sensitive to GMAT rise than other GCMs<sup>30</sup>, however, it provides an appropriate and important platform to test the tipping-point hypothesis (Supplementary Information). If the most sensitive of GCMs to greenhouse gas forcing does not illustrate tipping-point behaviour, we would not expect such behaviour in other, less sensitive models.

The finding that RILEs in model outcomes result from increased volatility of an ice cover that is progressively thinning because of warming temperatures—rather than tipping-point behaviour—is consistent with recently observed summer sea-ice declines. The sea-ice loss between September 2006 and September 2007, which was roughly equal to the entire loss of September ice extent between 1979 and 2006, encouraged speculation that a tipping point might have been crossed<sup>5</sup>. Yet, the 2008 and 2009 minima, although well below the long-term mean, were less severe than the record set in 2007<sup>6,10</sup>. Major losses of summer sea ice can thus occur, both in models and in observations, without pushing the sea ice past a tipping point into a permanent state of ice-free summers<sup>6,10,26</sup>. Instead of tipping-point behaviour, recent observations and model outcomes illustrate great natural variability superimposed on a secular warming-induced sea-ice decline. Controlling temperature increase, therefore, is the key to preserving sea-ice habitat.

We derived Bayesian network projections, informed by CCSM3 habitat projections, for polar bear populations in four ecoregions (Supplementary Fig. 8). With A1B habitat values, polar bears were most likely to disappear from the Seasonal Ice Ecoregion (SEA) and Polar Basin Divergent Ice Ecoregion (DIV) by mid-century, and to be substantially reduced in the Archipelago Ecoregion (ARC) and the Polar Basin Convergent Ice Ecoregion (CON). With MIT habitat values, extinction probabilities were much lower in all ecoregions



**Figure 5 | Greenhouse gas mitigation and best possible wildlife management could allow polar bears to persist throughout current range.** Bayesian network outcomes with habitat inputs from the MIT scenario are shown for the last decade of the twenty-first century. When temperature rise is kept at or below the MIT scenario and when on-the-ground management of harvest, bear–human interactions, oil and gas activities etc. is maximized (influence no. 2), extinction is not the most probable outcome in any of the four ecoregions.

(Fig. 4). Contrary to the A1B case, when greenhouse gas mitigation was combined with best on-the-ground management practices (for example, controlling hunting and other interactions with humans) extinction was not the most probable outcome in any ecoregion, and future population sizes in the CON and ARC could be equivalent to or even larger than at present (Fig. 5). Greenhouse gas mitigation that keeps GMAT rise below 1.25 °C combined with traditional wildlife management could, it seems, maintain polar bear numbers at sustainable although lower-than-present levels throughout the century. (Supplementary Information).

## METHODS SUMMARY

**Relationships between temperature and habitat.** We evaluated relationships between GMAT change and four habitat variables important to polar bear foraging success: resource-selection-function-based optimal habitat<sup>1</sup>; the temporal and spatial extent of sea ice over shallow continental shelf waters<sup>2,14,17</sup>; and the distance ice retreated from the continental shelf. GMAT change was calculated as the difference between the mean temperature of 1980–1999 (13.67 °C), and the future temperatures projected by CCSM3 under the different greenhouse gas scenarios. **Effects of habitat alteration on polar bears.** We projected the effects of habitat alteration on polar bear persistence with a Bayesian network model<sup>3</sup> modified to include inputs from other subject matter experts. Our Bayesian network model incorporated changes in four habitat variables projected for each of four ecoregions (Supplementary Fig. 8), with four greenhouse gas scenarios. The Bayesian

network model also was informed by the broad range of other currently available information including: potential anthropogenic stressors; the established links between reduced physical stature and survival and declining sea-ice availability among polar bears in parts of their range<sup>2,15–17</sup>; qualitative information indicating that similar processes are underway in parts of the polar bear range where quantitative data are not yet available; the fact that polar bears ultimately are dependent on the sea ice<sup>13,14</sup> for consistent foraging success; and knowledge that if greenhouse-gas-induced warming continues to increase, essential polar bear sea-ice habitats ultimately will disappear<sup>13</sup>.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 30 March; accepted 8 November 2010.**

- Durner, G. M. *et al.* Predicting 21st century polar bear habitat distribution from global climate models. *Ecol. Monogr.* **79**, 25–58 (2009).
- Regehr, E. V., Hunter, C. M., Caswell, H., Amstrup, S. C. & Stirling, I. Survival and breeding of polar bears in the southern Beaufort Sea in relation to sea ice. *J. Anim. Ecol.* **79**, 117–127 (2010).
- Amstrup, S. C., Marcot, B. G. & Douglas, D. C. in *Arctic Sea Ice Decline: Observations, Projections, Mechanisms, and Implications* (eds DeWeaver, E. T., Bitz, C. M. & Tremblay, L.-B.) 213–268 (American Geophysical Union, 2008).
- Lenton, T. M. *et al.* Tipping elements in the Earth's climate system. *Proc. Natl Acad. Sci. USA* **105**, 1786–1793 (2008).
- Kerr, R. A. Is battered Arctic sea ice down for the count? *Science* **318**, 33–34 (2007).
- Notz, D. The future of ice sheets and sea ice: between reversible retreat and unstoppable loss. *Proc. Natl Acad. Sci. USA* **106**, 20590–20595 (2009).
- Collins, W. D. *et al.* The Community Climate System Model version 3. *J. Clim.* **19**, 2122–2143 (2006).
- Lindsay, R. W. & Zhang, J. The thinning of Arctic sea ice, 1988–2003: have we passed a tipping point? *J. Clim.* **18**, 4879–4894 (2005).
- Holland, M. M., Bitz, C. M. & Tremblay, B. Future abrupt reductions in the summer Arctic sea ice. *Geophys. Res. Lett.* **33**, L23503 (2006).
- National Snow and Ice Data Center. Arctic sea ice extent remains low. *NSIDC Notes* 69, ([http://nsidc.org/pubs/notes/69/Notes\\_69\\_web.pdf](http://nsidc.org/pubs/notes/69/Notes_69_web.pdf)) (2009).
- Grosell, M. & Walsh, P. J. Sentinel species and animal models of human health. *Oceanography (Wash. D.C.)* **19**, 127–133 (2006).
- Washington, W. M. *et al.* How much climate change can be avoided by mitigation? *Geophys. Res. Lett.* **36**, L08703 (2009).
- Amstrup, S. C. *et al.* Rebuttal of “Polar bear population forecasts: a public-policy audit”. *Interfaces* **39**, 353–369 (2009).
- Thiemann, G. W., Derocher, A. E. & Stirling, I. Polar bear *Ursus maritimus* conservation in Canada: an ecological basis for identifying designatable units. *Oryx* **42**, 504–515 (2008).
- Stirling, I., Lunn, N. J. & Iacozza, J. Long-term trends in the population ecology of polar bears in western Hudson Bay in relation to climatic change. *Arctic* **52**, 294–306 (1999).
- Rode, K. D., Amstrup, S. C. & Regehr, E. V. Reduced body size and cub recruitment in polar bears associated with sea ice decline. *Ecol. Appl.* **20**, 768–782 (2010).
- Hunter, C. M. *et al.* Climate change threatens polar bear populations: a stochastic demographic analysis. *Ecology* **91**, 2883–2897 (2010).
- Nakićenović, N. *et al.* *Emissions Scenarios. A Special Report of Working Group III of the Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, 2000).
- Sokolov, A. P. *et al.* Probabilistic forecast for 21st century climate based on uncertainties in emissions (without policy) and climate parameters. *J. Clim.* **22**, 5175–5204 (2009).
- Ramanathan, V. & Feng, Y. On avoiding dangerous anthropogenic interference with the climate system: formidable challenges ahead. *Proc. Natl. Acad. Sci.* **105**, 14245–14250 (2008).
- Kizzia, T. Alaska polar bears called doomed. *Anchorage Daily News* A1 (September 8 2007).
- Wigley, T. M. L. The climate change commitment. *Science* **307**, 1766–1769 (2005).
- U.S. Climate Change Science Program. *Scenarios of Greenhouse Gas Emissions and Atmospheric Concentrations*. Sub-report 2.1a of Synthesis and Assessment Product 2.1 (Department of Energy, Office of Biological & Environmental Research, Washington DC, 2007).
- Hansen, J., Sato, M., Ruedy, R., Lacis, A. & Oinas, V. Global warming in the twenty-first century: an alternative scenario. *Proc. Natl Acad. Sci. USA* **97**, 9875–9880 (2000).
- Winton, M. in *Arctic Sea Ice Decline: Observations, Projections, Mechanisms, and Implications* (eds DeWeaver, E. T., Bitz, C. M. & Tremblay, L.-B.) 111–132 (American Geophysical Union, 2008).
- Eisenman, I. & Wettlaufer, J. S. Nonlinear threshold behavior during the loss of Arctic sea ice. *Proc. Natl Acad. Sci. USA* **106**, 28–32 (2009).
- Lawrence, D. M., Slater, A. G., Thomas, R. A., Holland, M. M. & Deser, C. Accelerated Arctic land warming and permafrost degradation during rapid sea ice loss. *Geophys. Res. Lett.* **35**, L11506 (2008).
- Holland, M. H., Bitz, C. M., Tremblay, L.-B. & Bailey, D. A. in *Arctic Sea Ice Decline: Observations, Projections, Mechanisms, and Implications* (eds DeWeaver, E. T., Bitz, C. M. & Tremblay, L.-B.) 133–150 (American Geophysical Union, 2008).
- Deser, C. & Teng, H. in *Arctic Sea Ice Decline: Observations, Projections, Mechanisms, and Implications* (eds DeWeaver, E. T., Bitz, C. M. & Tremblay, L.-B.) 7–26 (American Geophysical Union, 2008).
- Stroeve, J., Holland, M. M., Meier, W., Scambos, T. & Serreze, M. Arctic sea ice decline: faster than forecast. *Geophys. Res. Lett.* **34**, L09501 (2007).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** Principal funding for this project was provided by the USGS. B.G.M. acknowledges support from the USDA Forest Service, Pacific Northwest Research Station. E.T.D. acknowledges the support of the Office of Science (BER), US Department of Energy, under grant ER64735 to the University of Maryland. E.D.'s work also was supported by the National Science Foundation (NSF) during his employment there. The findings reported here, however, are not endorsed by and do not necessarily reflect the views of the NSF. CCSM3 simulations were performed using computing resources provided by the National Center for Atmospheric Research and the Earth Simulator in Japan. D.A.B. was supported under a grant from the NSF Office of Polar Programs, award number 0908675. M. Holland provided comments regarding model design and analysis. We acknowledge the Program for Climate Model Diagnosis and Intercomparison and the World Climate Research Programme's Working Group on Coupled Modeling for their roles in making available the Coupled Model Intercomparison Project phase 3 multi-model data set (support of this data set is provided by the Office of Science, US Department of Energy). We thank W. Washington and L. Buja for running the AS and providing us with the output from the CCSP integrations. We thank D. Vongraven and S. Vavrus for comments on earlier versions of this manuscript, and we thank N. Lunn and L. Peacock for providing the peer reviews necessary for the beta version of our Bayesian network model. Any use of trade names is for descriptive purposes only and does not represent endorsement by the US government.

**Author Contributions** S.C.A. conceived the project, assembled the team, and led writing. E.T.D. helped refine the project and analysed habitat/GMAT and RIGEs. D.C.D. staged sea-ice data and did the spatial analysis related to sea-ice metrics. B.G.M. conducted Bayesian network model runs and compiled outcomes. G.M.D. led development of the resource selection function approach to habitat analysis. C.M.B. proposed and helped interpret the 2020 CO<sub>2</sub> stabilization experiments. D.A.B. set up and ran the climate model simulations. E.T.D., C.M.B. and D.A.B. led interpretation of GCM outcomes. S.C.A., B.G.M. and D.C.D. interpreted biological outcomes. E.T.D. and D.C.D. developed all graphics. All authors contributed to writing and responding to review comments.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to S.C.A. ([samstrup@pbears.org](mailto:samstrup@pbears.org)).



## METHODS

**GCM and scenarios.** We used five emissions scenarios in our CCSM3 experiments. The Y2K scenario fixes atmospheric greenhouse gas concentrations at year 2000 levels<sup>22</sup>. The CCSP450<sup>23</sup> scenario keeps end of century total anthropogenic radiative forcing below  $3.4 \text{ W m}^{-2}$ , whereas the AS<sup>24</sup> does not allow anthropogenic radiative forcing to exceed  $1.5 \text{ W m}^{-2}$  above year 2000 levels. In A1B and B1,  $\text{CO}_2$  rises to 689 p.p.m. and 537 p.p.m. by 2100 (using the CCSM3 concentration values, see Supplementary Fig. 1 and Supplementary Table 1). Greenhouse gas concentrations for the CCSP450<sup>23</sup> and SRES<sup>18</sup> scenarios were calculated from the emissions specified for these scenarios with the Model for the Assessment of Greenhouse Gas Induced Climate Change<sup>22</sup>, a globally averaged gas-cycle/climate model. See ref. 12 for discussion of the CCSP greenhouse gas concentrations, and ref. 31 for details of the SRES<sup>18</sup> integrations. Greenhouse gas concentrations used in the AS are in supporting table 2 found at <http://www.pnas.org/content/101/46/16109/suppl/DC1>.

We obtained eight realizations each for A1B and B1, four realizations each for CCSP450 and Y2K, and one realization of the AS. Because net radiative forcing in the AS and CCSP450 were similar (Supplementary Fig. 2), and because global temperature change (Supplementary Fig. 3) and change in sea-ice extent (Supplementary Fig. 4) projected by the single AS run were very similar to members of the 4-run ensemble of CCSP450, we combined the single AS run with the 4 CCSP450 runs to create a 5-run mitigation ensemble (MIT). This left us with 4 forcing ensembles with which to compare the projected effects on the future welfare of polar bears: A1B, B1, Y2K and MIT.

GMAT change was calculated as the difference between the annual mean temperature of 1980–1999 ( $13.67^\circ\text{C}$ ), and the future temperatures projected by CCSM3 under the different greenhouse gas scenarios we examined. We derived the 1980–1999 mean from 8 CCSM3 model runs incorporating greenhouse gas increases observed through the twentieth century (20C3M ensemble)<sup>32</sup>.

**Ecoregions.** We evaluated how mitigation might affect polar bears occupying four Arctic ecoregions defined by temporal and spatial differences in observed ice melt, freeze, advection, bathymetry, proximity to land, and polar bear responses to those patterns (Supplementary Fig. 8). Each ecoregion is large, composed of several recognized subdivisions of the global polar bear population<sup>3,33</sup>, and not entirely homogeneous. Nonetheless, they offer useful subdivisions of the worldwide polar bear distribution because areas within each tend to be more similar than they are to portions of other ecoregions.

The SEA includes Hudson Bay, Foxe Basin, Baffin Bay and Davis Strait. There, sea ice melts entirely in summer and the ~7,500 bears occurring there are forced ashore for extended periods during which they are largely food deprived. The ARC—the channels between the Canadian Arctic Islands—is presently home to ~5,000 bears and is characterized by heavy sea ice, much of which is present year round. The polar basin (the portion of the Arctic Ocean centred on the North Pole and ringed by the continental shelves of Eurasia, North America, Greenland and the Canadian Archipelago; Supplementary Fig. 8) was divided into a DIV, including the Southern Beaufort, Chukchi, East Siberian–Laptev, Kara and Barents Seas, and a CON including the east Greenland Sea, the continental shelf areas adjacent to northern Greenland and the Queen Elizabeth Islands, and the northern Beaufort Sea. Extensive formation of annual sea ice occurs in the DIV where ~8,500 bears currently occur. That ice typically is advected towards the central polar basin, out of the polar basin through Fram Strait, or against the CON. The CON is currently home to ~2,400 polar bears. Differences among ecoregions acknowledge that global warming effects on sea-ice habitats have different starting points<sup>14</sup> and that the nature of sea-ice changes is likely to be different.

**Habitat metrics.** We examined the relationship between GMAT change and four habitat variables known to be important to polar bears. First, we adopted the resource selection function (RSF) approach previously described<sup>1</sup> to convert GCM projections of sea-ice extent to projections of optimal polar bear habitat. RSFs are quantitative expressions of the habitats animals choose to utilize, relative to available habitats and resources<sup>34</sup>. Sea-ice concentrations for the observational period were estimated from monthly passive-microwave (PMW) satellite imagery<sup>35</sup>. Choices polar bears made from among available habitats were determined from 1985–1995 satellite radiolocations<sup>1</sup>. Optimal habitat was defined as any mapped pixel with an RSF value in the upper 20% of the seasonally averaged (1985–1995) RSF scores, and could be expressed as the sum of qualifying mapped pixels over any period of interest. We assessed changes in habitat availability by comparing annual sums of optimal habitat among projected time periods<sup>1</sup>.

Estimates of optimal habitat were limited to the polar basin because only there did we have access to the radio-tracking data necessary to build RSF models. The importance of sea ice over continental shelves, however, is widely recognized as an important component of polar bear habitat<sup>1,14</sup>. Therefore, we derived a second habitat variable we called ‘total shelf-ice habitat’ from both observed and projected Arctic-wide sea-ice concentration maps. Total shelf-ice habitat was defined as the

aerial cover ( $\text{km}^2$ ) of all pixels with  $\geq 50\%$  ice concentration that were mapped over the continental shelves ( $< 300 \text{ m}$  depth). Waters with less than 50% ice cover were denoted ice-free because available data indicate that areas with sea-ice coverage  $< 50\%$  may not be preferred<sup>1,15</sup>. Unlike optimal habitat, total shelf-ice habitat could be calculated in all ecoregions and therefore provided a means of quantifying projected changes in habitat availability throughout the range of polar bears. We compared shelf-ice habitat expressed as the annual 12-month sum of sea-ice extent over the continental shelves in each ecoregion. Because SEA and ARC are almost entirely continental shelf area, the total shelf-ice habitat in those ecoregions equated to the total annual area (sum of 12 months) of  $\geq 50\%$  concentration sea ice.

The third habitat variable, one of the most important variables representing seasonal changes in habitat available to polar bears<sup>2,15–17,36</sup>, was calculated as the change from present in the number of months that ice was projected to be absent (ice-free months) from the continental shelves. An ice-free month occurred in an ecoregion when  $< 50\%$  of the shelf area was covered by sea ice of  $\geq 50\%$  concentration. Outside the polar basin this variable represented simply the ice-free season because the SEA and ARC are composed almost entirely of continental shelf.

Recognizing that the magnitude of the separation of the sea ice from preferred foraging areas also might be important, we calculated a fourth habitat variable as the change in average distance from the continental shelf to the ice pack during the month of minimum ice extent (shelf-to-ice distance). Shelf-to-ice distance was calculated, for the month of minimum ice extent, as the mean distance from every shelf pixel in either of the polar basin ecoregions to the nearest ice-covered pixel ( $> 50\%$  concentration) in the main body of perennial ice. We did not calculate shelf-to-ice distance in SEA and ARC because they are almost entirely comprised of continental shelf.

We plotted GMAT change against these habitat features to evaluate potential nonlinearities in the relationships. Figure 2 and Supplementary Figs 5 and 6 illustrate annual mean GMAT values (x-axis) and corresponding habitat values (y-axis) for each year of each simulation (small dots). Each scenario is shown in a different colour. Large connected dots in each plot are centred on the means, over all years, of the annual GMAT values and values of the habitat-related variables, where GMAT lies within  $0.25^\circ\text{C}$  bins centred on  $0.25^\circ\text{C}$ ,  $0.5^\circ\text{C}$ ,  $0.75^\circ\text{C}$ , etc., for all simulations performed for each scenario. Large dots are not in exact vertical alignment because the means of the GMAT values in each bin differ among scenarios.

**Bayesian network model.** The effects of future habitat alteration on probabilities of future polar bear persistence were projected with a beta version<sup>37</sup> of the Bayesian network model used previously<sup>3</sup>. The beta model was reviewed by two other polar bear experts and modified accordingly. Some conditional probabilities were modified to incorporate reviewers’ suggestions and observations noted since building the original model. The beta model includes a finer division of bins for sea-ice habitat variables, but upper and lower bounds were retained to ensure that the range of possible entries in conditional probability tables was consistent with the assignments in ref. 3. The final structure (nodes and links) of the beta model is nearly identical to that of the alpha model<sup>3</sup>.

Our beta model incorporated changes in the four habitat variables projected under different greenhouse gas scenarios. We calculated the average per cent of future changes, from the 2001–2010 decade, in annual optimal and shelf-ice habitat. Changes in the number of ice-free months and the shelf-to-ice distance were expressed as the average increases (months of ice absence and kilometres of ice retreat) at each decade.

The Bayesian network model also was informed by the broad range of other currently available information including: potential anthropogenic stressors; the established links between reduced physical stature and survival and declining sea-ice availability among polar bears in parts of their range<sup>2,15–17,36</sup>; available qualitative information indicating that similar processes are underway in parts of the polar bear range where quantitative data are not yet available; the fact that polar bears ultimately are dependent on the sea ice<sup>13,14</sup> for consistent foraging success; and that if greenhouse-gas-induced warming continues to increase, essential polar bear sea-ice habitats ultimately will disappear<sup>13</sup>. These additional factors were incorporated into the model as ordinal or qualitative categories or as background with which conditional probability tables were parameterized. The beta model incorporated 4 greenhouse gas scenarios and was applied to each of the four ecoregions at four future decadal time periods: 2020–2029, 2045–2054, 2070–2079 and 2090–2099. At each time period, states of these variables could represent a condition similar to present, better than present, or worse than present (see tables 3 and 4 in ref. 3).

31. Meehl, G. A. *et al.* Climate change projections for the 21st century and climate change commitment in the CCSM3. *J. Clim.* **19**, 2597–2616 (2006).
32. Meehl, G. A. *et al.* Combinations of natural and anthropogenic forcings in twentieth-century climate. *J. Clim.* **17**, 3721–3727 (2004).
33. International Union for Conservation of Nature. *Polar Bears: Proceedings of the Fourteenth Working Meeting of the IUCN/SSC Polar Bear Specialists Group* (eds Aars, J., Lunn, N. J. & Derocher, A. E.) occasional paper no. 32, 189 (IUCN, 2006).

34. Manly, B. F. J., McDonald, L. L., Thomas, D. L., McDonald, T. L. & Erickson, W. P. *Resource Selection by Animals: Statistical Design and Analysis for Field Studies* (Kluwer Academic, 2002).
35. Cavalieri, D., Parkinson, C., Gloersen, P. & Zwally, H. J. Sea ice concentrations from Nimbus-7 SMMR and DMSP SSM/I passive microwave data (Digital media, National Snow and Ice Data Center, Boulder, Colorado, 1996, updated, 2006).
36. Molnár, P. K., Derocher, A. E., Thiemann, G. W. & Lewis, M. A. Predicting survival, reproduction and abundance of polar bears under climate change. *Biol. Conserv.* **143**, 1612–1622 (2010).
37. Marcot, B. G., Steventon, J. D., Sutherland, G. D. & McCann, R. K. Guidelines for developing and updating Bayesian belief networks applied to ecological modeling and conservation. *Can. J. For. Res.* **36**, 3063–3074 (2006).

# Intercalation of a new tier of transcription regulation into an ancient circuit

Lauren N. Booth<sup>1</sup>, Brian B. Tuch<sup>1†</sup> & Alexander D. Johnson<sup>1</sup>

Changes in gene regulatory networks are a major source of evolutionary novelty<sup>1–3</sup>. Here we describe a specific type of network rewiring event, one that intercalates a new level of transcriptional control into an ancient circuit. We deduce that, over evolutionary time, the direct ancestral connections between a regulator and its target genes were broken and replaced by indirect connections, preserving the overall logic of the ancestral circuit but producing a new behaviour. The example was uncovered through a series of experiments in three ascomycete yeasts: the bakers' yeast *Saccharomyces cerevisiae*, the dairy yeast *Kluyveromyces lactis* and the human pathogen *Candida albicans*. All three species have three cell types: two mating-competent cell forms ( $\alpha$  and  $a$ ) and the product of their mating ( $a/\alpha$ ), which is mating-incompetent. In the ancestral mating circuit, two homeodomain proteins, Mata1 and Mata2, form a heterodimer that directly represses four genes that are expressed only in  $a$  and  $\alpha$  cells and are required for mating<sup>4–6</sup>. In a relatively recent ancestor of *K. lactis*, a reorganization occurred. The Mata1–Mata2 heterodimer represses the same four genes (known as the core haploid-specific genes) but now does so indirectly through an intermediate regulatory protein, Rme1. The overall logic of the ancestral circuit is preserved (haploid-specific genes ON in  $a$  and  $\alpha$  cells and OFF in  $a/\alpha$  cells), but a new phenotype was produced by the rewiring: unlike *S. cerevisiae* and *C. albicans*, *K. lactis* integrates nutritional signals, by means of Rme1, into the decision of whether or not to mate.

In *S. cerevisiae*, *K. lactis* and *C. albicans*, three cell types ( $a$ ,  $\alpha$  and  $a/\alpha$ ) are specified by transcriptional regulators (sequence-specific DNA-binding proteins) encoded at the mating type locus. An important part of this cell-type-specific circuit is the regulation of the haploid-specific genes (hsgs), a group of genes that are expressed in  $a$  and  $\alpha$  cells but not in  $a/\alpha$  cells<sup>7</sup>. The full sets of hsgs were previously identified in *S. cerevisiae*<sup>4</sup> and *C. albicans* (ref. 5, and B.B.T., Q. M. Mitrovich, F. M. De La Vega, C. K. Monighetti and A.D.J., unpublished observations) but not in the related species *K. lactis*. To examine the evolution of this portion of the mating circuit, we identified the genes in the *K. lactis* hsg regulon and compared them to those in *S. cerevisiae* and *C. albicans*. By profiling the expression patterns of wild-type  $a$ ,  $\alpha$  and  $a/\alpha$  *K. lactis* cells genome-wide, we identified 12 genes that are clear hsgs under the conditions tested (Fig. 1a), two of which—*RME1* (referred to previously as *MTS1*) and *STE4*—were previously identified as hsgs in *K. lactis*<sup>8</sup>. Comparison of all the hsgs in the three species revealed a substantial level of turnover in the regulon; in other words, an hsg in one species is not necessarily an hsg in the other two (Fig. 1b). However, an ancestral core of four hsgs (*GPA1*, *STE4*, *STE18* and *FAR1*) share a common expression pattern in all three species. The first three genes encode the heterotrimeric G protein that, in the presence of mating pheromone, activates a downstream mitogen-activated protein kinase (MAPK) cascade required for mating<sup>9–11</sup>. Far1 lies further downstream in this pathway and mediates two responses needed as a prelude to mating, cell cycle arrest<sup>12</sup> and the formation of mating projections<sup>13,14</sup>.

In *S. cerevisiae* and *C. albicans*  $a/\alpha$  cells, all four genes of the hsg core regulon are directly repressed by the transcription regulator  $a1$ – $\alpha2$  (refs 4, 6), a heterodimer encoded by one gene at the MAT $\alpha$  locus and one gene at the MAT $a$  locus. To determine whether this was also true in *K. lactis*, genome-wide chromatin immunoprecipitations (ChIP-chip) of  $a1$  and  $\alpha2$  were performed in *K. lactis*  $a/\alpha$  cells. In total, the upstream regions of 14 genes were observed to be occupied by both  $a1$  and  $\alpha2$ , including the *RME1* gene (Fig. 1c), which is also  $a1$ – $\alpha2$  regulated in *S. cerevisiae*.  $a1$  and  $\alpha2$  ChIP peaks were not observed at the promoters of any of the four core hsgs in *K. lactis*, indicating that, unlike in *S. cerevisiae* and *C. albicans*, these genes are not directly regulated by  $a1$ – $\alpha2$ .

To confirm the absence of direct  $a1$ – $\alpha2$  regulation at *K. lactis* hsgs, we identified the  $a1$ – $\alpha2$  recognition motif in *K. lactis* from the ChIP data, using a *de novo* motif-finding program<sup>15</sup>. The highest-scoring motif was similar to the  $a1$ – $\alpha2$  motifs previously identified in *S. cerevisiae*<sup>16</sup> and *C. albicans*<sup>17</sup> (Fig. 1d). Indeed, the *S. cerevisiae* motif is efficiently recognized by the *C. albicans*  $a1$ – $\alpha2$  protein<sup>17</sup>, confirming that key features of this sequence have remained largely unchanged in the three species. We searched the regions 2 kilobases upstream of each *K. lactis* core hsg for the *K. lactis*  $a1$ – $\alpha2$  motif but did not find significant matches, confirming the absence of direct  $a1$ – $\alpha2$  regulation of these genes. (Whereas the  $a1$ – $\alpha2$  site upstream of *RME1* had a log<sub>10</sub>-odds score of 4.98, the best matches at the core hsgs ranged from –0.70 to 0.93.) These results indicate that although the ancestral core hsg expression pattern is conserved in *K. lactis*, the mechanism of the regulation has changed.

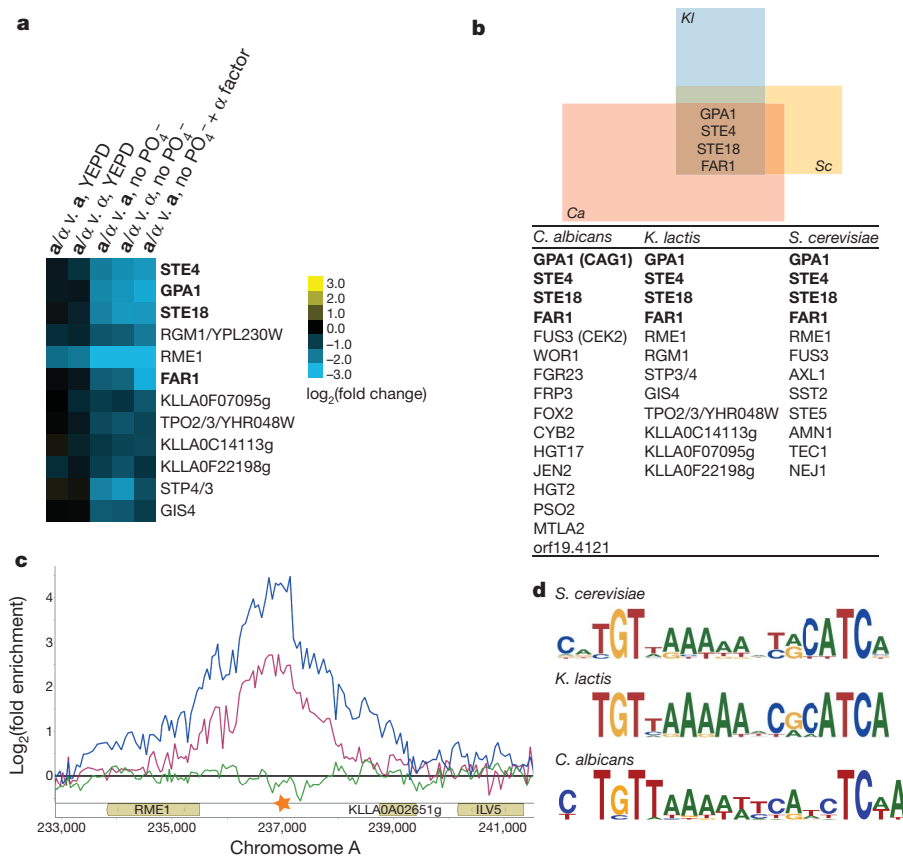
To understand how the *K. lactis* hsgs are cell-type regulated we searched the upstream regions of the 12 genes identified as hsgs by expression array (Fig. 1a) for *cis*-regulatory motifs<sup>15</sup>. The second-highest ranking motif (the top-ranking motif was a repeat sequence) was found in 11 out of 12 of the promoters (Fig. 2a) and was similar to the *S. cerevisiae* Rme1 motif (*K. lactis* consensus, GAACCNMAA; *S. cerevisiae* consensus, GAACCTCAA<sup>18,19</sup>). This motif is also similar to, although longer than, the *K. lactis* Rme1 motif derived previously<sup>20</sup>. The Rme1 motif is absent from *S. cerevisiae* and *C. albicans* hsg promoters.

In *S. cerevisiae*, Rme1 was initially identified as a repressor of meiosis and sporulation<sup>21,22</sup>, and was later shown to act as a transcriptional activator of other genes<sup>19</sup>. In *K. lactis*, Rme1 was shown to regulate mating-type interconversion<sup>20</sup> (the switching of  $a$  and  $\alpha$  cells to the opposite cell-type by means of DNA rearrangement). We speculated that Rme1 was co-opted in the *K. lactis* lineage to positively regulate the core hsgs.

To test this hypothesis, we knocked out the *RME1* gene in *K. lactis*  $a$  cells and examined the gene expression profile by microarray. We found that 20 genes were downregulated in the knockout strain (Fig. 2b), including all four of the core hsgs ( $P = 2 \times 10^{-10}$ , hypergeometric distribution). We also observed a set of genes that was upregulated in the absence of Rme1 (Supplementary Fig. 1), including a significant number of genes orthologous to *S. cerevisiae* sporulation genes (Gene Ontology (GO): 0043934,  $n = 14$ ,  $P = 10^{-14}$ , hypergeometric

<sup>1</sup>Department of Microbiology and Immunology and Department of Biochemistry and Biophysics, University of California, San Francisco, California 94158, USA. †Present address: Genome Analysis Unit, Amgen, South San Francisco, California 94080, USA.





**Figure 1 | The core hsgs are not directly regulated by  $\alpha 1$ – $\alpha 2$  in *K. lactis*.**

**a**, The expression profiles of the set of 12 hsgs identified in *K. lactis*. Note that phosphate starvation induces expression of the hsgs and is required to identify these genes. For example, when starved for phosphate, the heterotrimeric G protein genes are expressed in  $\alpha$  and  $\alpha$  cells at levels about fivefold higher than in  $\alpha/\alpha$  cells. **b**, A comparison of hsgs defined by transcriptional profiling in *S. cerevisiae* (Sc)<sup>4</sup>, *K. lactis* (Kl) (panel **a**) and *C. albicans* (Ca) (ref. 5, and B.B.T., Q. M. Mitrovich, F. M. De La Vega, C. K. Monighetti and A.D.J., unpublished observations) shows a conserved subset of hsgs (GPA1, STE4, STE18, FAR1),

distribution) indicating that the function of Rme1 in regulating meiosis is shared with *S. cerevisiae*. Thus, whereas some of the targets of Rme1 have remained the same as in the common ancestor of *S. cerevisiae* and *K. lactis*, Rme1 has gained new targets in the *K. lactis* lineage, including the core hsgs.

To test whether Rme1 directly regulates the core hsgs in *K. lactis*, we performed a genome-wide ChIP of Rme1 in  $\alpha$  cells. ChIP peaks were observed at the promoters of the four core hsgs (Fig. 2c) and were centred over the Rme1 motifs (Fig. 2c). Thus, in the *K. lactis* lineage, Rme1 was gained as a direct activator of the core hsgs by the acquisition of Rme1 *cis*-regulatory sequences at all four genes. We note that Rme1 is not the only regulator of the *K. lactis* hsgs; for example, STE18 is repressed by Sir2 (ref. 23).

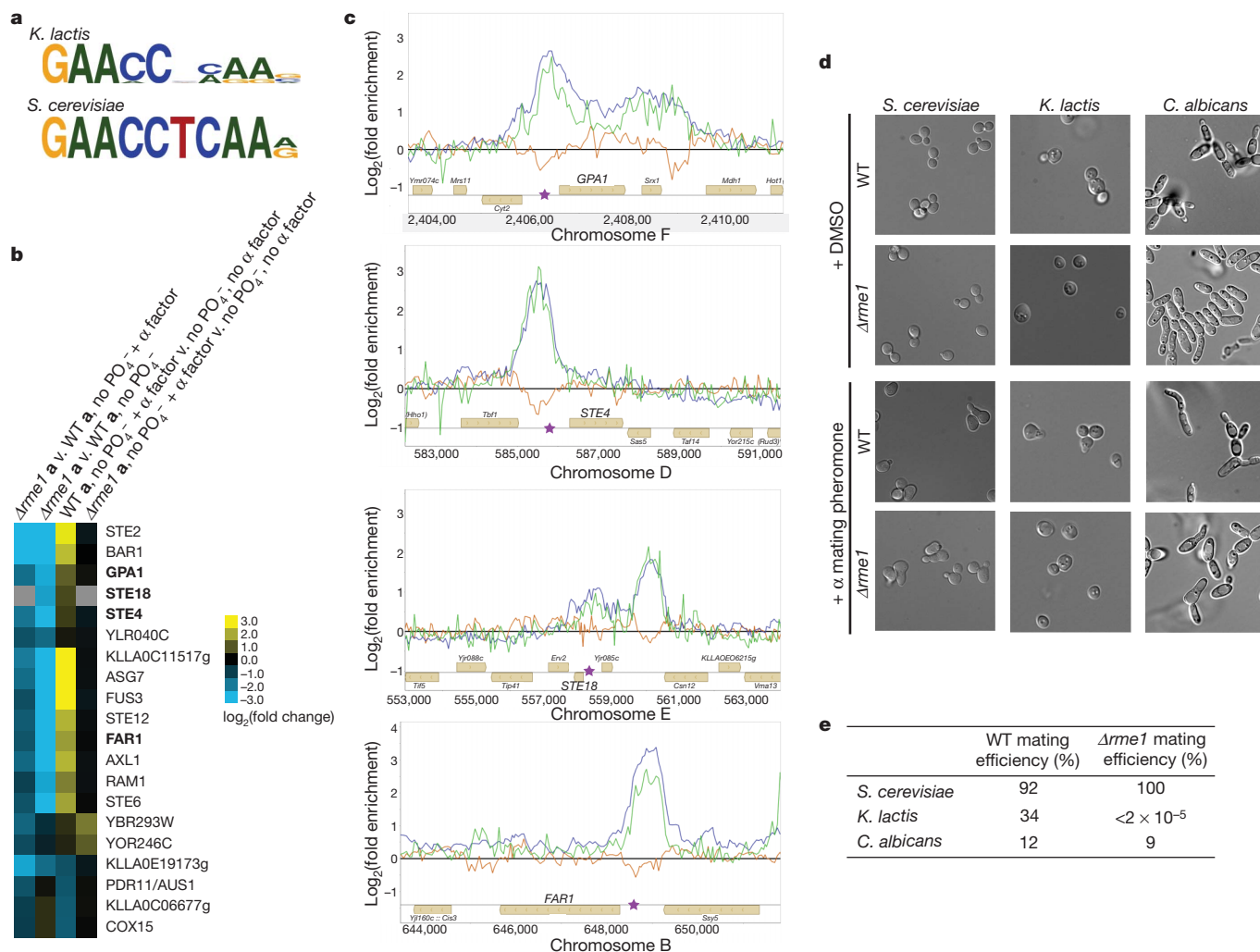
We next tested the biological role of Rme1 in mating in *K. lactis*, *S. cerevisiae* and *C. albicans* by comparing wild-type and RME1 knockout  $\alpha$  cells. In response to  $\alpha$  pheromone,  $\alpha$  cells form mating projections (polarized growth towards the source of pheromone). When *S. cerevisiae* and *C. albicans* wild-type and *Arme1*  $\alpha$  cells were exposed to  $\alpha$  mating pheromone, both strains formed mating projections normally (Fig. 2d). In contrast, whereas *K. lactis* wild-type  $\alpha$  cells produced mating projections in response to pheromone, *Arme1*  $\alpha$  cells did not, indicating that this biological response was dependent on Rme1 (Fig. 2d). As a second test of the role of Rme1, we examined mating directly using a quantitative mating assay. No difference was observed between the mating efficiencies of wild-type  $\alpha$  cells and those of *Arme1*  $\alpha$  cells for *S. cerevisiae* and *C. albicans* (Fig. 2e). In contrast,

which we refer to as the core hsgs (bold in **a** and **b**). **c**, ChIP enrichment profiles from experiments using haemagglutinin (HA)-tagged MAT $\alpha 1$   $\alpha/\alpha$  cells (magenta), HA-tagged MAT $\alpha 2$   $\alpha/\alpha$  cells (blue) and, as a control, untagged  $\alpha/\alpha$  cells (green). The ChIP enrichment was determined by hybridization to a tiling microarray. The location of the  $\alpha 1$ – $\alpha 2$  motif in the RME1 promoter is indicated by the orange star. The genes (tan boxes) are all transcribed in the reverse direction. Data were visualized with MochiView<sup>30</sup>. **d**, The *K. lactis*  $\alpha 1$ – $\alpha 2$  motif determined from the ChIP-chip data. For comparison, the *S. cerevisiae* and *C. albicans* motifs (derived from published ChIP data<sup>4,6</sup>) are also shown.

the *K. lactis* *Arme1*  $\alpha$  cell mating efficiency was decreased, relative to the wild type, by a factor of at least  $10^6$  (Fig. 2e). Thus, the ability to mate is critically dependent on Rme1—but only in *K. lactis*.

Unlike *S. cerevisiae* and *C. albicans*, *K. lactis* requires a starvation signal to mate<sup>24</sup> and to respond to pheromone<sup>25</sup>. Although several different types of starvation signal can prime *K. lactis* to respond to pheromone<sup>24,25</sup>, we found that phosphate starvation is particularly potent, and it was used in subsequent experiments. Our expression profiling experiments (Fig. 1a) revealed that *K. lactis* requires starvation to express most of its mating genes. RME1 was also highly induced (24-fold) by phosphate starvation (Fig. 1a). We note that *S. cerevisiae* RME1 transcript levels also increase tenfold under starvation conditions<sup>18</sup>, suggesting that regulation of RME1 by starvation may be ancestral to *S. cerevisiae* and *K. lactis*.

We next investigated in greater detail how the starvation signal is incorporated in the *K. lactis* mating regulatory circuit. The simplest model consistent with the data presented so far is that starvation upregulates RME1, which in turn activates transcription of the hsgs. A prediction of this model is that ectopic expression of RME1 in *K. lactis* should override the requirement for starvation in expressing the hsgs. We created an  $\alpha$  strain overexpressing RME1 to levels that were within tenfold of the level in starved wild-type cells (using the Kl LAC4 promoter) and found that overexpression of RME1 (in rich YEP-galactose medium) was sufficient to induce expression of the heterotrimeric G protein subunits (Fig. 3a). Overexpression of RME1 is also sufficient to allow *K. lactis* to form mating projections



**Figure 2 | RME1 is a direct activator of hsg expression and is required for *K. lactis* mating.** **a**, The *K. lactis* Rme1 motif found by a *de novo* search<sup>15</sup> of the 12 *Kl* hsgs and the *S. cerevisiae* motif derived from two experimentally characterized binding sites<sup>18,19</sup>. **b**, The set of 19 genes repressed twofold or greater relative to wild type when *RME1* is absent and the cells are phosphate-starved. In bold are the core hsgs. **c**, Rme1 is a direct regulator of the core hsgs. ChIP of Rme1 was performed in *K. lactis* c-Myc-tagged *RME1*  $\alpha$  cells (blue and green lines, two biological replicates) and untagged, control  $\alpha$  cells (orange line). The immunoprecipitated DNA was hybridized to a tiling microarray. The genes (tan boxes) above the line are transcribed in the forward direction and those below are transcribed in the reverse direction. The location of the *Kl* Rme1 motif is indicated by a purple star. **d**, Rme1 is required only in *K. lactis* to

respond to mating pheromone. Wild-type or *RME1* knockout  $\alpha$  cells were exposed to either  $\alpha$  mating pheromone or a mock treatment of dimethylsulphoxide (DMSO). Mating projections form readily when both wild-type and  $\Delta$ rme1 *S. cerevisiae* and *C. albicans* cells are exposed to mating pheromone. Only the *K. lactis*  $\Delta$ rme1  $\alpha$  cells were unable to respond to the presence of  $\alpha$  mating pheromone. **e**, Rme1 is required for mating in *K. lactis* but not in *S. cerevisiae* nor *C. albicans*. Quantitative mating assays were performed by mating wild-type or  $\Delta$ rme1  $\alpha$  cells to wild-type  $\alpha$  cells. In *S. cerevisiae* and *C. albicans* the percentage of  $\alpha$  cells that was able to mate is similar for wild-type and  $\Delta$ rme1. *K. lactis*  $\Delta$ rme1  $\alpha$  cells are mating incompetent; no mating products were isolated from the  $\Delta$ rme1  $\alpha$   $\times$  wild-type  $\alpha$  mating.

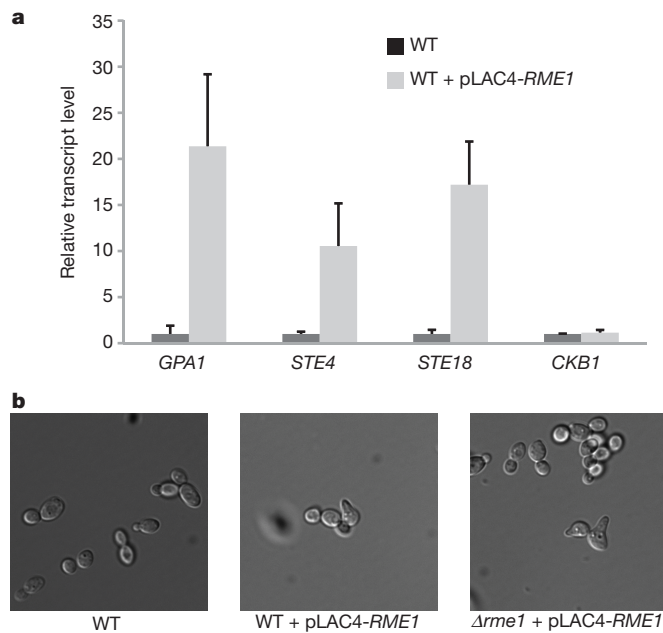
in response to pheromone in rich medium (Fig. 3b). These results strongly support the model by showing that upregulation of *RME1* is sufficient to cause biologically relevant upregulation of the heterotrimeric G proteins.

Thus, the rewiring of the *K. lactis* hsg circuit (summarized in Fig. 4) resulted in a new network configuration and a novel phenotype, relative to the ancestor. Our results suggest a possible evolutionary path for this rewiring. In the ancestor of all three yeasts, the hsgs were directly repressed by  $\alpha 1$ – $\alpha 2$ . Either in an ancestor to *S. cerevisiae* and *K. lactis* or independently in each lineage, *RME1* was brought under nutritional regulation. Finally, in the *K. lactis* lineage alone, two steps occurred: the hsgs lost the *cis*-regulatory sequences for  $\alpha 1$ – $\alpha 2$  and gained the *cis*-regulatory sequences for Rme1. As described in Supplementary Fig. 2, it is possible to determine more precisely when the rewiring of the core hsgs occurred. We can infer that direct  $\alpha 1$ – $\alpha 2$  regulation of the core hsgs was probably lost several times in the ascomycete lineage, and that the *K. lactis* form of regulation of the hsgs probably arose after *K. lactis*

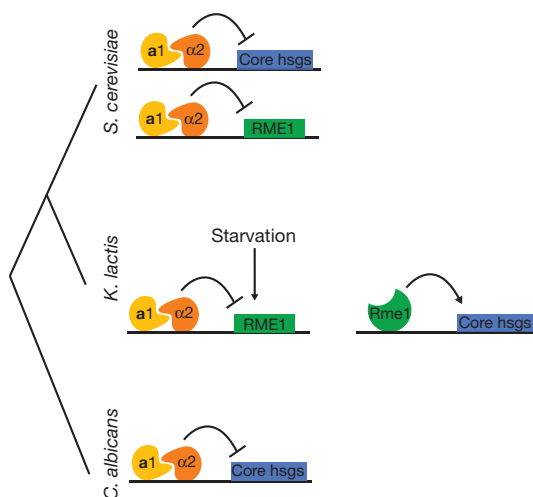
and the closely related species *L. kluyveri* branched from their common ancestor.

Although we do not know whether acquisition of the *K. lactis* form of regulation was adaptive, this type of regulation makes logical sense given that the primary mode of growth of *K. lactis* is as a haploid<sup>26</sup>. The formation of spores is a strategy employed by many yeasts to survive harsh environments. For starvation to give rise to spores, *K. lactis* would first have to mate (to form the sporulation-competent  $\alpha/\alpha$  cell type), thus rationalizing the link between starvation and mating. In contrast, *S. cerevisiae* in the wild is typically at least diploid<sup>27</sup> and forms spores directly in response to starvation. Thus, the coupling of mating and starvation makes conceptual sense for *K. lactis* in comparison with *S. cerevisiae*.

We have described a case in which a new tier of regulation has been intercalated into an ancient transcription circuit consisting of a regulator (a homeodomain heterodimer) and a set of target genes. This change involved breaking the original connections between the regulator and its



**Figure 3 | Overexpression of *RME1* is sufficient for hsg expression in the absence of nutrient starvation.** **a**, In the overexpression strain (pLAC4-*RME1*), *RME1* transcription is induced by galactose-containing medium, a condition that does not cause expression of the heterotrimeric G proteins or pheromone response in wild-type (WT) cells. A strain using the empty pLAC4 vector was used as a control. The transcripts were measured relative to *ACT1* transcript levels by RT-quantitative PCR (means and s.d.,  $n = 3$ ). In the absence of a starvation signal the hsgs, but not *CKB1* (a non-hsg control), are upregulated when *RME1* is overexpressed. **b**, *RME1* overexpression allows cells to respond to mating pheromone in the absence of a starvation signal. *K. lactis* **a** cells that contained only the endogenous *RME1* copy and an empty pLAC4 vector (WT), the endogenous copy of *RME1* and *RME1* driven by the pLAC4 promoter (WT + pLAC4-*RME1*) or only *RME1* driven by the pLAC4 promoter ( $\Delta rme1$  + pLAC4-*RME1*) were grown in YEP-galactose and exposed to  $\alpha$  mating pheromone. Wild-type cells were unable to form mating projections in the absence of a starvation signal, but both strains overexpressing *RME1* (pLAC4-*RME1*) formed mating projections in the absence of a starvation signal.



**Figure 4 | A simplified model for the evolution of regulation of core hsgs in three yeasts.** In all three species the core hsgs are repressed by  $\alpha 1$ – $\alpha 2$ ; thus, they are ON in **a** and  $\alpha$  cells and OFF in **a**/ $\alpha$  cells. In *S. cerevisiae* and *C. albicans* the repression is direct ( $\alpha 1$ – $\alpha 2$  binds to the promoters of these genes), but in *K. lactis* it is indirect, through Rme1. The circuit rewiring in the *K. lactis* lineage has resulted in a new mating behaviour; this species is able to mate only when starved. We show that this behaviour is due to the intercalation of Rme1, which is upregulated by starvation in *K. lactis*.

target genes and replacing them with a more complex type of hierarchy (Fig. 4). Intercalation may be a common way in which regulatory circuits evolve. This type of ‘intercalary evolution’ was first proposed<sup>28</sup> to account for a common origin of eyes. In a wide variety of species, the transcription regulator Pax6 lies at the top of the eye development hierarchy, and rhodopsins occupy the bottom. According to the proposal, different types of eye arose from evolutionary intercalation of a variety of regulatory and structural genes within this simple, deeply conserved, regulatory relationship. The change we describe here is less complex and provides a concrete example of evolutionary intercalation, one that is responsible for an important feature of modern mating behaviour in *K. lactis*. It has been known for decades that *K. lactis* (unlike its relatives *S. cerevisiae* and *C. albicans*) requires starvation to mate<sup>24</sup>, and we have shown that this behaviour is due to the new configuration of the *K. lactis* mating circuit.

## METHODS SUMMARY

**Gene expression array.** RNA was isolated by hot phenol extraction and reverse transcribed, and the resulting complementary DNAs were coupled to Cy5. A pooled mixture of the cDNAs was coupled to Cy3 and used as a reference. Labelled cDNAs were hybridized to Agilent arrays for visualization.

**ChIP-chip.** ChIP experiments were performed as described previously<sup>29</sup>, with minor modifications.

**Pheromone response assays.** Exponential-phase cultures were exposed to 13-mer  $\alpha$ -mating pheromone, and the formation of mating projections was monitored by microscopy.

**Quantitative mating assay.** **a** and  $\alpha$  cultures were grown independently to exponential phase and then mixed together with  $\alpha$  cells in fivefold excess under mating conditions. The mating products were selected for on medium that either **a** and **a**/ $\alpha$  cells or only **a**/ $\alpha$  cells could grow on, and efficiencies were calculated as efficiency = (**a**/ $\alpha$  colonies)/(**a** and **a**/ $\alpha$  colonies).

**RT-quantitative PCR.** Cultures were grown to exponential phase in YEP-galactose medium, and RNA was isolated by extraction with hot phenol. RNA was reverse transcribed and the cDNAs were quantified by quantitative PCR.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 27 April; accepted 6 October 2010.**

- Carroll, S. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134**, 25–36 (2008).
- Davidson, E. H. & Erwin, D. H. Gene regulatory networks and the evolution of animal body plans. *Science* **311**, 796–800 (2006).
- Wray, G. A. The evolutionary significance of cis-regulatory mutations. *Nature Rev. Genet.* **8**, 206–216 (2007).
- Galgoczy, D. J. et al. Genomic dissection of the cell-type-specification circuit in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* **101**, 18069–18074 (2004).
- Tsong, A. E., Miller, M. G., Raisner, R. M. & Johnson, A. D. Evolution of a combinatorial transcriptional circuit: a case study in yeasts. *Cell* **115**, 389–399 (2003).
- Srikantha, T. et al. TOS9 regulates white-opaque switching in *Candida albicans*. *Eukaryot. Cell* **5**, 1674–1687 (2006).
- Herskowitz, I. A regulatory hierarchy for cell specialization in yeast. *Nature* **342**, 749–757 (1989).
- Barsoum, E., Sjöstrand, J. O. O. & Åström, S. U. Ume6 is required for the MAT $\alpha$ /MAT $\alpha$ -cellular identity and transcriptional silencing in *Kluyveromyces lactis*. *Genetics* **184**, 999–1011 (2010).
- Herskowitz, I. MAP kinase pathways in yeast: for mating and more. *Cell* **80**, 187–197 (1995).
- Dignard, D., André, D. & Whiteway, M. Heterotrimeric G-protein subunit function in *Candida albicans*: both the  $\alpha$  and  $\beta$  subunits of the pheromone response G protein are required for mating. *Eukaryot. Cell* **7**, 1591–1599 (2008).
- Coria, R. et al. The pheromone response pathway of *Kluyveromyces lactis*. *FEM. Yeast Res.* **6**, 336–344 (2006).
- Chang, F. & Herskowitz, I. Identification of a gene necessary for cell cycle arrest by a negative growth factor of yeast: FAR1 is an inhibitor of a G1 cyclin, CLN2. *Cell* **63**, 999–1011 (1990).
- Butty, A. C., Pryciak, P. M., Huang, L. S., Herskowitz, I. & Peter, M. The role of Far1p in linking the heterotrimeric G protein to polarity establishment proteins during yeast mating. *Science* **282**, 1511–1516 (1998).
- Côte, P. & Whiteway, M. The role of *Candida albicans* FAR1 in regulation of pheromone-mediated mating, gene expression and cell cycle arrest. *Mol. Microbiol.* **68**, 392–404 (2008).
- Bailey, T. L. & Elkan, C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **2**, 28–36 (1994).



16. Goutte, C. & Johnson, A. D. Recognition of a DNA operator by a dimer composed of two different homeodomain proteins. *EMBO J.* **13**, 1434–1442 (1994).
17. Hull, C. M. & Johnson, A. D. Identification of a mating type-like locus in the asexual pathogenic yeast *Candida albicans*. *Science* **285**, 1271–1275 (1999).
18. Covitz, P. A. & Mitchell, A. P. Repression by the yeast meiotic inhibitor RME1. *Genes Dev.* **7**, 1598–1608 (1993).
19. Toone, W. M. *et al.* Rme1, a negative regulator of meiosis, is also a positive activator of G1 cyclin gene expression. *EMBO J.* **14**, 5824–5832 (1995).
20. Barsoum, E., Martinez, P. & Astrom, S. U.  $\alpha 3$ , a transposable element that promotes host sexual reproduction. *Genes Dev.* **24**, 33–44 (2010).
21. Kassir, Y. & Simchen, G. Regulation of mating and meiosis in yeast by the mating-type region. *Genetics* **82**, 187–206 (1976).
22. Mitchell, A. P. & Herskowitz, I. Activation of meiosis and sporulation by repression of the RME1 product in yeast. *Nature* **319**, 738–742 (1986).
23. Hickman, M. A. & Rusche, L. N. The Sir2-Sum1 complex represses transcription using both promoter-specific and long-range mechanisms to regulate cell identity and sexual cycle in the yeast *Kluyveromyces lactis*. *PLoS Genet.* **5**, e1000710 (2009).
24. Herman, A. Interspecies sex-specific growth responses in *Kluyveromyces*. *Antonie van Leeuwenhoek* **36**, 421–425 (1970).
25. Tuch, B. B., Galgoczy, D. J., Hernday, A. D., Li, H. & Johnson, A. D. The evolution of combinatorial gene regulation in fungi. *PLoS Biol.* **6**, e38 (2008).
26. Kurtzman, C. P. & Fell, J. W. *The Yeasts: A Taxonomic Study* 4th edn (Elsevier, 2000).
27. Ezov, T. K. *et al.* Molecular-genetic biodiversity in a natural population of the yeast *Saccharomyces cerevisiae* from 'Evolution Canyon': microsatellite polymorphism, ploidy and controversial sexual status. *Genetics* **174**, 1455–1468 (2006).
28. Gehring, W. J. & Ikeo, K. *Pax 6*: mastering eye morphogenesis and eye evolution. *Trends Genet.* **15**, 371–377 (1999).
29. Nobile, C. J. *et al.* Biofilm matrix regulation by *Candida albicans* Zap1. *PLoS Biol.* **7**, e1000133 (2009).
30. Homann, O. R. & Johnson, A. D. MochiView: versatile software for genome browsing and DNA motif analysis. *BMC Biol.* **8**, 49 (2010).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank Q. Mitrovich, O. Homann, A. Hernday, M. Miller, C. Cain, T. Sorrells and H. Madhani for helpful discussions and technical contributions; and S. Åström for generously providing the *K. lactis* strains used in this study. The *S. cerevisiae* strains were a gift from the H. Madhani and J. Li laboratories. The work was funded by grant RO1 GM037049 from the National Institutes of Health. L.N.B. is a National Science Foundation Graduate Research Fellow.

**Author Contributions** L.N.B. performed all experiments. L.N.B. and B.B.T. analysed data. L.N.B., B.B.T. and A.D.J. designed the study and wrote the paper.

**Author Information** The gene expression array data have been deposited in the NCBI Gene Expression Omnibus (GEO) under accession number GSE24874. For the ChIP-chip data the accession number is GSE25209. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to A.D.J. ([ajohnson@cgl.ucsf.edu](mailto:ajohnson@cgl.ucsf.edu)).

## METHODS

**Medium.** Details of the medium used in the experiments presented here can be found in ref. 31. The recipe for the phosphate starvation medium can be found in ref. 25.

**Strains and strain construction.** The strains used in this study can be found in Supplementary Information. *S. cerevisiae* strains are S288C background and *C. albicans* strains are SC5314 background.

Gene disruption cassettes for knockouts and taggings in *K. lactis* and *L. kluyveri* were generated by fusion PCR<sup>32</sup> using the primers listed in Supplementary Information. Fusion PCRs were performed in a 50- $\mu$ l reaction containing 0.5  $\mu$ l ExTaq (Takara Bio Inc.), 0.25 mM dNTPs, 0.2  $\mu$ M each primer and about 25 ng of template. The reactions were incubated as follows: 94 °C for 3 min; 35 cycles of 94 °C for 30 s, 50–55 °C (depending on primer) for 30 s and 72 °C for 1 min per kilobase; and 72 °C for 5 min. The first round of PCR consisted of three reactions that amplified the flanking homologous sequence from *K. lactis* genomic DNA with primers 1 and 3 or 4 and 6, and amplified the markers from the appropriate plasmids with primers 2 and 5. The URA3 marker was amplified from YEp24, the TRP1 marker from YEplac112, the c-Myc tagging cassette from pFA6a-13Myc-kanMX6 (ref. 33) and the 3  $\times$  HA tagging cassette from pYMN-20 (ref. 34). The products were purified with the QIAquick PCR Purification Kit (Qiagen). The second round of amplification (the fusion round) used 1  $\mu$ l of each purified flank PCR product and 2  $\mu$ l of the purified marker PCR product. This product was purified with the QIAquick PCR Purification Kit.

The purified fusion PCR products were transformed into *K. lactis* and *L. kluyveri* by electroporation<sup>35,36</sup>. Transformants were confirmed to be correct by colony PCR with the check primers listed in Supplementary Information. Tagged genes were also verified by sequencing.

To avoid mating-type switching in *K. lactis* and to increase the efficiency of tagging genes in the MAT loci, we created strains in which the silent MAT loci (HML $\alpha$  and HMR $\alpha$ ) were knocked out. SAY509 and SAY572 were transformed with gene disruption cassettes for HMR $\alpha$  (yLB10b) and HML $\alpha$  (yLB11b), respectively. To generate strains with both silent cassettes deleted, these strains were mated as described<sup>35</sup> for 2 days, and diploids were selected for by URA/TRP drop-out SCD medium. The resulting strain (yLB12d) was sporulated on pre-SPO plates for 2 days, and haploids with both silent cassettes deleted were selected for on URA/TRP drop-out medium. The mating types were determined by colony PCR using the MAT, MAT $\alpha$  and MAT $\alpha$  check primers. These strains (yLB13a and yLB14) were mated to generate an a/ $\alpha$  strain lacking silent MAT loci (yLB15b).

To create N-terminally 3  $\times$  HA-tagged MAT $\alpha$ 1 and MAT $\alpha$ 2 for ChIP experiments, yLB13a and yLB14 were transformed with the gene disruption cassettes created with the HA primers and the MAT $\alpha$ 1-tag or MAT $\alpha$ 2-tag primers. yLB55 was mated with yLB14 to create yLB58, the a/ $\alpha$  strain used in the MAT $\alpha$ 1 ChIP experiment. yLB56a was mated with yLB13a to create yLB57a, the a/ $\alpha$  strain used in the MAT $\alpha$ 2 ChIP experiment.

RME1 was knocked-out (yLB21a) and tagged (yLB54b) in SAY572.

yDG957 is the mated product of SAY509 and SAY572.

yLB33a1 was created by sorbose-selecting TF028X, as described in ref. 37. AH136a1 was sorbose-selected from SN87, and HIS1 and LEU2 were added back as described in ref. 38.

pLAC4 (described below) was transformed into SAY572 after digestion with SacII (New England Biolabs) to generate yLB61b. pLAC4-RME1 (described below) was transformed into SAY 572 and yLB21a to create yLB64a and yLB65, respectively.

pLAC4 was created by modifying pKLAC1 (New England Biolabs) as follows. The pKLAC1 vector was cut with HindIII and XhoI and gel-purified with the QIAquick Gel Extraction Kit to remove most of the  $\alpha$ -mating pheromone secretory domain. The vector was dephosphorylated with APEX heat-labile dephosphatase (Epicentre), in accordance with the kit's instructions. Primers BamHI add and BamHI add, reverse complement were hybridized to each other at a concentration of 40  $\mu$ M in 1  $\times$  T4 PNK buffer (New England Biolabs) and 40 mM additional NaCl by incubation at 94 °C for 2 min and slow cooling to 10 °C at a rate of 0.1 °C s<sup>-1</sup>. The hybridized primers created sticky ends for HindIII and XhoI cut sites. The hybridized primers were phosphorylated by adding ATP to a final concentration of 1 mM and T4 PNK kinase (New England Biolabs) to a final concentration of 200 U ml<sup>-1</sup> and incubating them at 37 °C for 10 min. The cut, dephosphorylated vector was ligated to the hybridized primers by using Fast-Link Ligase (Epicentre) at a 1:5 molar ratio and transformed into DH5 $\alpha$  cells. The acetimidase marker in the plasmid was then replaced with kanMX6. kanMX6 was amplified from pFA6a-13myc-kanMX6 (ref. 33) by using the KAN primers. The plasmid and kanMX6 marker were cut with BsrGI and XmaI (New England Biolabs) and ligated and transformed as described above. This plasmid was used as an empty vector control (referred to as pLAC4) in the RME1 overexpression experiments.

pLAC4-RME1 was generated by amplifying RME1 from *K. lactis* genomic DNA with the primers RME1 + BglII and RME1 + NotI. The RME1 gene and the pLB12 plasmid were cut with BglII and NotI (New England Biolabs) and ligated and transformed as described above.

**Orthologous gene set mapping.** In this study we used the orthologous gene sets defined previously<sup>25</sup>.

**Gene expression arrays.** Arrays were designed with OligoArray (v2.1.3)<sup>39</sup>. The reference sequence used was downloaded from the NCBI genome project for *Kluyveromyces lactis* NRRL Y-1140, records NC006037 to NC006042. The predicted messenger RNA sequences were used. The sequences of two shorter open reading frame (ORF)-coding transcripts, MFA1 and AGA2, that were not annotated in NCBI at the time were added to the reference. The following parameters were used: maximum number of oligonucleotides to design per input sequence, 3; size range, 60 to 60; maximum distance between the 5' end of the oligonucleotide and the 3' end of the input sequence, 1,500; minimum distance between the 5' ends of two adjacent oligonucleotides, 69; T<sub>m</sub> range, 75 to 97 °C; GC range, 15.0 to 65.0; threshold to reject secondary structures, 65.0; threshold to start to consider cross-hybridizations, 65; sequence to avoid in the oligonucleotide: GGGGG;CCCCC;TTTTT;AAAAA.

The arrays were printed by Agilent, using the 4  $\times$  44K format.

*K. lactis* strains were grown in either rich medium (YEPD) to an attenuation ( $D_{600}$ ) of 0.9 or in phosphate starvation medium, with or without  $\alpha$ -mating pheromone, as described previously<sup>25</sup>. The 50-ml cultures were centrifuged for 5 min at 3,700g, and the pellet was resuspended in 10 ml of 1  $\times$  TE and centrifuged again. The supernatant was removed and pellets were frozen in liquid nitrogen and stored at -80 °C.

RNA was isolated and reverse transcribed as described previously<sup>38</sup> with the exception that the RNA isolation protocol was scaled to 50-ml cultures and that SuperScript II (Invitrogen) was used. Additionally, reverse transcription reactions were performed on either individual samples or on RNA in which an equivalent amount of each RNA sample was pooled.

For the individual samples, 1.3  $\mu$ g of cDNA was dried and resuspended in 5  $\mu$ l of 0.1 M sodium bicarbonate. For the pooled samples, 5.9  $\mu$ g of cDNA was dried and resuspended in 22.5  $\mu$ l of 0.1 M sodium bicarbonate. An equivalent volume of Cy3 (pool) or Cy5 (individual) dye (Amersham) was added (dyes were resuspended in 60  $\mu$ l of DMSO) and the reaction was incubated in the dark at 65 °C for 20 min. Labelled cDNAs were purified with a Clean and Concentrator -5 kit (Zymo Research).

Equal amounts of the Cy3-labelled and Cy5-labelled cDNA were hybridized overnight to the array, as described in the Agilent protocol. After hybridization, the arrays were washed as specified by Agilent. Arrays were scanned at 5  $\mu$ m, averaging two lines, with an Axon GenePix 4000A scanner. Arrays were gridded with GenePix Pro v5.1. Global Lowess normalization analysis was performed for each array with a Goulphar script<sup>40</sup> (R Foundation for Statistical Computing). Normalized data were collapsed first by averaging the result for all duplicate probes and finally by taking the median of the probes for each ORF. Data were transformed as described for each experiment. Microarray data were clustered with Cluster version 3.0 (ref. 41) and visualized with Java TreeView v1.1.3 (ref. 42).

**ChIP-chip.** *K. lactis* was grown in either YEPD to  $D_{600}$  = 0.4 or in phosphate starvation medium as described previously<sup>25</sup>. The ChIP, DNA amplification, labelling and hybridization were performed as described previously<sup>29</sup>. For the a1 and  $\alpha$ 2 ChIPs, 2  $\mu$ l of 5 mg ml<sup>-1</sup> mouse anti-HA antibody clone 12CA5 (Roche) were used. The *K. lactis* tiling arrays used have been described previously<sup>25</sup>. Peak finding was performed with the 'Extract peaks from Data Set(s)' utility of MochiView<sup>40</sup>. Peak extraction was applied independently to each normalized ChIP-chip data item. Peak finding significance thresholds were kept at their default values ( $P \leq 0.001$ ). For the a1 and  $\alpha$ 2 ChIPs, regions of overlap between the two ChIP-chips were determined and the overlapping chromosomal coordinates were extracted, yielding 22 regions. This set was filtered to remove peaks in the telomeres or those that fell entirely in an ORF. MEME was performed on the remaining 14 peaks.

**Pheromone response assay.** For *K. lactis* (SAY572 and yLB21a), the pheromone response assay was performed as described previously<sup>25</sup>, with the exception that the medium was not supplemented with additional leucine. For *C. albicans* (AH136a1 and yLB33a1), cells were grown in SCD medium at room temperature (23–25 °C) to  $D_{600}$  = 1.0 and exposed to pheromone (or mock 10% DMSO treatment) for 4 h as described in ref. 43. For *S. cerevisiae* (BY4674 and YMI768), the cells were grown in YEPD at 30 °C to  $D_{600}$  = 0.5.  $\alpha$ -Mating pheromone (5  $\mu$ M; Sigma-Aldrich) was added and the cells were grown at 30 °C for 90 min; the presence of mating projections was monitored by microscopy.

**Quantitative mating.** The strains and selective medium used for these experiments are listed in Supplementary Information.

Yeasts were grown in YEPD (*S. cerevisiae*, *K. lactis*) or SCD (*C. albicans*) at 30 °C (*S. cerevisiae*, *K. lactis*) or room temperature (*C. albicans*) to  $D_{600}$  = 0.8. a cells ( $2 \times 10^6$ ; about 200  $\mu$ l of  $D_{600}$  = 0.8) were mixed with  $10^7$   $\alpha$  cells (about 1 ml of

$D_{600} = 0.8$ ) in 10 ml of YEPD. The mixtures were filtered onto nitrocellulose (0.8  $\mu\text{m}$  pore size; Millipore) with a Millipore 1225 Vacuum Manifold. Filter discs were placed onto either YEPD + 55  $\mu\text{g ml}^{-1}$  adenine 2% agar plates (*S. cerevisiae*, *C. albicans*) or 2% malt extract, 3% agar plates (*K. lactis*). *S. cerevisiae* strains were allowed to mate for 5 h at 30 °C, *K. lactis* strains for 2 days at 30 °C, and *C. albicans* strains for 6 days at room temperature. The filter discs were then removed and placed in 5 ml of SD medium and the cells were dispersed by vortex-mixing. The cell suspensions were sonicated with a Branson Sonifier 450 at 30% power for 10 s. Between  $1/25$  and  $2.5 \times 10^{-6}$  of the cell suspension was plated onto selective medium (see Supplementary Information). *S. cerevisiae* and *K. lactis* cells were grown at 30 °C for 2 days, and *C. albicans* cells were grown at room temperature for 4 days. Mating efficiency was calculated with the limiting parental  $\alpha$  cells: efficiency =  $(a/\alpha)/(a + \alpha/\alpha)$ .

**RT-quantitative PCR.** *K. lactis* strains were grown in YEP-galactose to  $D_{600} = 0.8$  and centrifuged at 3,700g for 5 min. The pellets were washed with 1 ml of  $1 \times \text{TE}$  and centrifuged at 3,700g for 5 min; supernatant was removed and frozen in liquid nitrogen. RNA was isolated and reverse transcribed (using SuperScript II) as described previously<sup>38</sup>, with all volumes scaled appropriately. cDNAs were quantified with a Bio-Rad CFX96 Real Time machine in a standard 25- $\mu\text{l}$  reaction using Sybr green under standard conditions. The primers used are listed in Supplementary Information.

**Hypergeometric test.** The significance of the set of genes ( $n = 20$ ) down twofold or greater relative to wild type in a phosphate-starved *Rme1* knockout containing all four core hsgs was calculated with a hypergeometric test. The background set of genes ( $n = 4,769$ ) are defined as all genes that were detectable by our gene expression array.

**GO term analysis.** For the analysis of the genes that were upregulated in the absence of RME1 in *K. lactis*, we performed GO term analysis using the *S. cerevisiae* GO term finder on SGD (<http://www.yeastgenome.org/cgi-bin/GO/goTermFinder.pl>). Genes that were up fourfold or greater in an RME1 knockout versus wild type in the RME1 array and that had an orthologue in *S. cerevisiae* were compared with a background

set of genes defined as all *S. cerevisiae* genes orthologous to *K. lactis* genes that could be detected in our gene expression array.

31. Sherman, F. Getting started with yeast. *Methods Enzymol.* **350**, 3–41 (2002).
32. Wach, A. PCR-synthesis of marker cassettes with long flanking homology regions for gene disruptions in *S. cerevisiae*. *Yeast* **12**, 259–265 (1996).
33. Longtine, M. S. et al. Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* **14**, 953–961 (1998).
34. Janke, C. et al. A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. *Yeast* **21**, 947–962 (2004).
35. Wolf, K., Breunig, K. & Barth, G. *Non-Conventional Yeasts in Genetics, Biochemistry and Biotechnology* (Springer, 2003).
36. Gojkovic, Z., Jahnke, K., Schnackerz, K. D. & Piskur, J. PYD2 encodes 5,6-dihydropyrimidine amidohydrolase, which participates in a novel fungal catabolic pathway. *J. Mol. Biol.* **295**, 1073–1087 (2000).
37. Janbon, G., Sherman, F. & Rustchenko, E. Monosomy of a specific chromosome determines L-sorbose utilization: a novel regulatory mechanism in *Candida albicans*. *Proc. Natl Acad. Sci. USA* **95**, 5150–5155 (1998).
38. Mitrovich, Q. M., Tuch, B. B., Guthrie, C. & Johnson, A. D. Computational and experimental approaches double the number of known introns in the pathogenic yeast *Candida albicans*. *Genome Res.* **17**, 492–502 (2007).
39. Rouillard, J.-M., Zuker, M. & Gulari, E. OligoArray 2.0: design of oligonucleotide probes for DNA microarrays using a thermodynamic approach. *Nucleic Acids Res.* **31**, 3057–3062 (2003).
40. Lemoine, S., Combes, F., Servant, N. & Le Crom, S. Goulphar: rapid access and expertise for standard two-color microarray normalization methods. *BMC Bioinformatics* **7**, 467 (2006).
41. de Hoon, M. J., Imoto, S., Nolan, J. & Miyano, S. Open source clustering software. *Bioinformatics* **20**, 1453–1454 (2004).
42. Saldanha, A. J. Java Treeview—extensible visualization of microarray data. *Bioinformatics* **20**, 3246–3248 (2004).
43. Bennett, R. J., Uhl, M. A., Miller, M. G. & Johnson, A. D. Identification and characterization of a *Candida albicans* mating pheromone. *Mol. Cell. Biol.* **23**, 8189–8201 (2003).



# Noise correlations improve response fidelity and stimulus encoding

Jon Cafaro<sup>2</sup> & Fred Rieke<sup>1,2</sup>

Computation in the nervous system often relies on the integration of signals from parallel circuits with different functional properties. Correlated noise in these inputs can, in principle, have diverse and dramatic effects on the reliability of the resulting computations<sup>1–8</sup>. Such theoretical predictions have rarely been tested experimentally because of a scarcity of preparations that permit measurement of both the covariation of a neuron's input signals and the effect on a cell's output of manipulating such covariation. Here we introduce a method to measure covariation of the excitatory and inhibitory inputs a cell receives. This method revealed strong correlated noise in the inputs to two types of retinal ganglion cell. Eliminating correlated noise without changing other input properties substantially decreased the accuracy with which a cell's spike outputs encoded light inputs. Thus, covariation of excitatory and inhibitory inputs can be a critical determinant of the reliability of neural coding and computation.

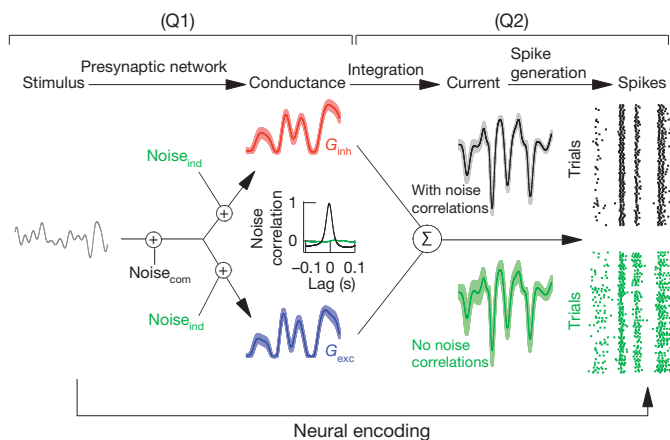
Differences in the properties of excitatory and inhibitory synaptic inputs to a target cell provide a key control of neural activity. Feed-forward inhibitory synaptic input is a ubiquitous example. A delay in inhibitory input relative to excitatory input, for example by an extra synaptic delay in the circuit providing inhibitory input, can limit response duration to the time window in which the target cell receives excitatory but not inhibitory input<sup>9</sup>. More generally, inhibitory input can cancel unwanted responses by arriving before or at the same time as excitatory input<sup>10–13</sup>. Theoretical work illustrates how the effectiveness of these computations depends on the strength of covariation between excitatory and inhibitory synaptic inputs<sup>8</sup>. Thus, although synaptic noise will always decrease the reliability of the neural response, strong noise correlations, unlike independent noise, could allow fluctuations in inhibitory synaptic input to cancel corresponding fluctuations in excitatory synaptic input<sup>5</sup> (Fig. 1). Such noise correlations can arise if noise within excitatory and inhibitory pathways originates from a common source (Fig. 1, left), for example in densely and randomly connected recurrent networks<sup>14</sup>. Noise cancellation in synaptic integration could in turn reduce trial-to-trial variability in a cell's spike output (Fig. 1, right).

The extent and impact of noise correlations depends on several network and cellular properties, including nonlinearities in synaptic transmission<sup>15</sup> or spike generation<sup>16</sup> that could decrease correlation strength. This dependence makes it difficult to predict the importance of noise correlations from modelling alone or from correlations measured in cell pairs. Work on the retina provides a rare opportunity to provide quantitative experimental information about how noise correlations affect the coding of physiologically relevant stimuli. Our goal was first to measure covariation of the excitatory and inhibitory synaptic inputs received by a retinal ganglion cell (Fig. 1, (Q1)) and then to test how these noise correlations affect the encoding of light stimuli in a cell's spike output (Fig. 1, (Q2)).

Quantifying the covariation of excitatory and inhibitory synaptic input requires measuring these two conductances simultaneously or near simultaneously. To do this, we rapidly alternated the ganglion cell voltage between the reversal potentials for excitatory and inhibitory

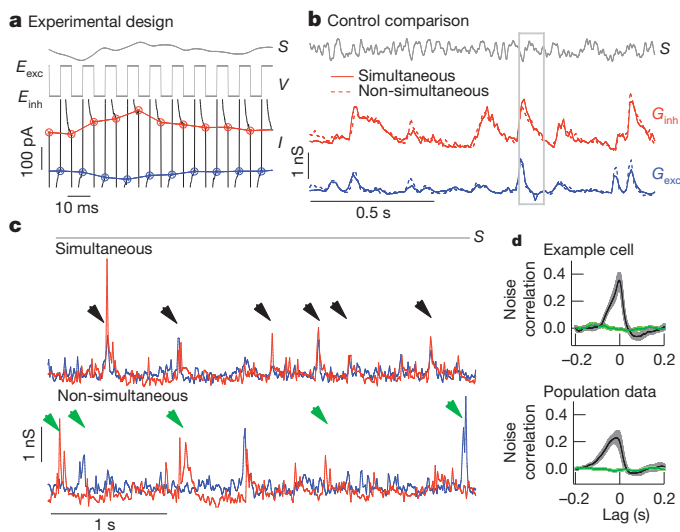
synaptic inputs, collecting a single sample of each input every 10 ms (Fig. 2a). Control experiments indicated that the voltage at the synaptic receptors had reached a near-constant value at these sampling times (Supplementary Fig. 1). This sampling rate is high in comparison with the 50–100 ms time course of a ganglion cell's response to light inputs. To check how well this procedure captured light-dependent changes in conductance, we compared the simultaneously measured conductances with those measured non-simultaneously when the voltage was held constant at the excitatory or inhibitory reversal potential. Mean excitatory and inhibitory conductances resulting from a repeated, modulated light input differed minimally (Fig. 2b). In 21 cells, the alternating voltage approach captured  $99.9 \pm 0.6\%$  of the power of the conductance signal and  $83 \pm 4\%$  of that of the conductance noise (mean  $\pm$  s.e.m.; see Methods). Thus, simultaneous conductance measurements capture most of the structure in the synaptic inputs a ganglion cell receives.

Simultaneous conductances measured during constant light input often exhibited spontaneous excitatory synaptic events accompanied in time by inhibitory synaptic events (Fig. 2c, top, black arrowheads). Such events in fact typically occurred together. Correlated noise events were rarely observed during non-simultaneously measured conductances (Fig. 2c, bottom). Correspondingly, the cross-correlation function for



**Figure 1 | Effects of noise correlations on the variability of synaptic current and spike output.** Neural encoding consists of three basic steps: a stimulus shapes excitatory (blue,  $G_{exc}$ ) and inhibitory (red,  $G_{inh}$ ) synaptic conductances; these conductances then shape synaptic currents; and the resulting currents control spike generation to produce a sequence of action potentials (spikes). Noise correlations will be strong if a common source dominates noise in excitatory and inhibitory pathways ( $Noise_{com}$ ) and minimal if the dominant noise source arises independently ( $Noise_{ind}$ ). Correlated (black traces) as opposed to uncorrelated (green traces) noise between excitatory and inhibitory conductances can lead to lower variability of both the synaptic current and the spike output (shaded regions around traces). Understanding this issue requires answering two questions. (Q1) How much do converging excitatory and inhibitory input covary? (Q2) What is the impact of such noise correlations on the neural output?

<sup>1</sup>Howard Hughes Medical Institute, University of Washington, Seattle, Washington 98195, USA. <sup>2</sup>Department of Physiology and Biophysics, University of Washington, Seattle, Washington 98195, USA.



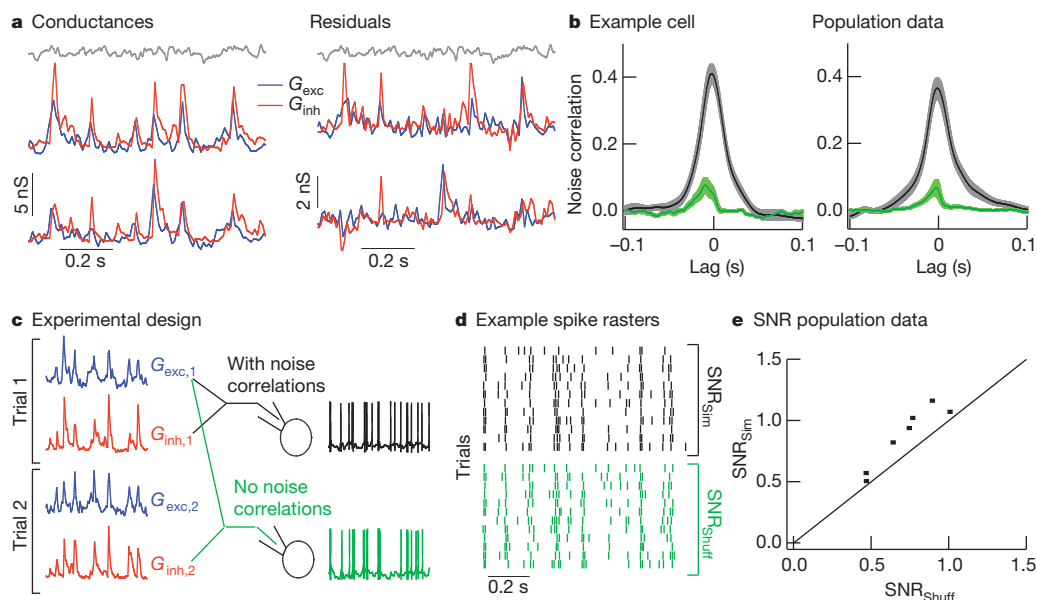
simultaneously measured excitatory and inhibitory conductances during constant light input showed considerable structure, unlike the cross-correlation for non-simultaneously measured conductances (single cell: Fig. 2d, top; population: Fig. 2d, bottom). Thus, simultaneous conductance recordings revealed correlations between converging synaptic inputs that were inaccessible from more conventional recordings.

To determine both the strength of noise correlations during modulated light input and their effect on a cell's spike output, we first considered midset ganglion cells, which comprise the majority of ganglion cells in the primate retina<sup>17</sup>. Midget ganglion cells receive delayed feed-forward synaptic inhibition, where the delay reflects an extra synapse in the circuit controlling inhibitory input. Thus, excitatory input comes directly from bipolar cells, whereas inhibitory input comes from amacrine cells that themselves receive input from bipolar cells<sup>18</sup>. Similar delayed feed-forward inhibition is a characteristic of many cortical

**Figure 2 | Near-simultaneous recording of excitatory and inhibitory synaptic input to an ON-OFF directionally selective ganglion cell.** **a**, Light stimulus (S) is presented while the voltage (V) of the cell alternates between the excitatory ( $E_{exc}$ ) and inhibitory ( $E_{inh}$ ) reversal potentials. Excitatory (blue) and inhibitory (red) synaptic currents (I) are sampled at the end of each voltage step. **b**, Conductances derived from measured currents (Methods) and averaged across multiple repeats of the same stimulus (S). Simultaneously measured conductances (solid lines) closely match those (dashed lines) measured non-simultaneously with the voltage held fixed at the reversal potentials for excitatory or inhibitory input (both excitatory and inhibitory correlations are  $0.91 \pm 0.01$  (mean  $\pm$  s.e.m.), 21 cells). **a** is an enlarged view of the boxed region of **b**. **c**, Top: section of simultaneously recorded conductances during constant light input shows correlated excitatory and inhibitory spontaneous events (black arrowheads). Bottom: non-simultaneously recorded conductances also show spontaneous events (green arrowheads), but they are rarely correlated. Records have been resampled at 50 Hz for comparison with the top conductances. **d**, Top: cross-correlation (mean  $\pm$  s.e.m., 10 trials) of excitatory and inhibitory conductances in an example cell during simultaneous (black) and non-simultaneous (green) recording. Bottom, cross-correlation for all recorded cells (mean  $\pm$  s.e.m., 6 cells).

circuits, including hippocampus, cerebellum, barrel cortex and auditory cortex<sup>9,11,12,19–21</sup>. We simultaneously recorded excitatory and inhibitory synaptic inputs during a full-field modulated light stimulus (Fig. 3a, left) and estimated variability in the synaptic responses by subtracting the average synaptic input from each individual trial (Fig. 3a, right). The peak correlation strength of the resulting residuals ranged from 0.15 to 0.5 (Fig. 3b, black traces). Noise correlations in the interleaved non-simultaneous conductances were substantially smaller (Fig. 3b, green traces). Slow drift in the light response accounted for the remaining noise correlations in the non-simultaneous conductances (Supplementary Fig. 2).

The alternating-voltage technique could produce artefactual noise correlations by overshooting the appropriate reversal potentials for excitatory or inhibitory synaptic inputs. For example, holding at a voltage positive relative to the excitatory reversal potential could cause



**Figure 3 | Strength and impact of noise correlations in synaptic inputs to primate midget ganglion cells.** **a**, Left: two trials of simultaneously recorded conductances during modulated light input (grey). Right: residual conductances (trials from left with mean subtracted), which estimate noise in each trial. **b**, Left: cross-correlation (mean  $\pm$  s.e.m., 12 trials) of excitatory and inhibitory residual conductances in an example cell during simultaneous (black) and non-simultaneous (green) recording. Right: cross-correlation for all recorded cells (mean  $\pm$  s.e.m., 15 cells). **c**, Logic of dynamic-clamp

experiments using simultaneously or shuffled simultaneous conductances in place of synaptic input. **d**, Example spike trains from 12 dynamic-clamp trials of simultaneous conductances (black) or their shuffled counterparts (green). SNR, signal-to-noise ratio. **e**, Signal-to-noise ratio of spike trains generated from simultaneous conductances versus that of spike trains generated from shuffled conductances (dots). The signal-to-noise ratio for simultaneous conductances was  $1.22 \pm 0.04$  times higher than that for shuffled conductances (mean  $\pm$  s.e.m., 7 cells,  $P = 0.0015$ ).

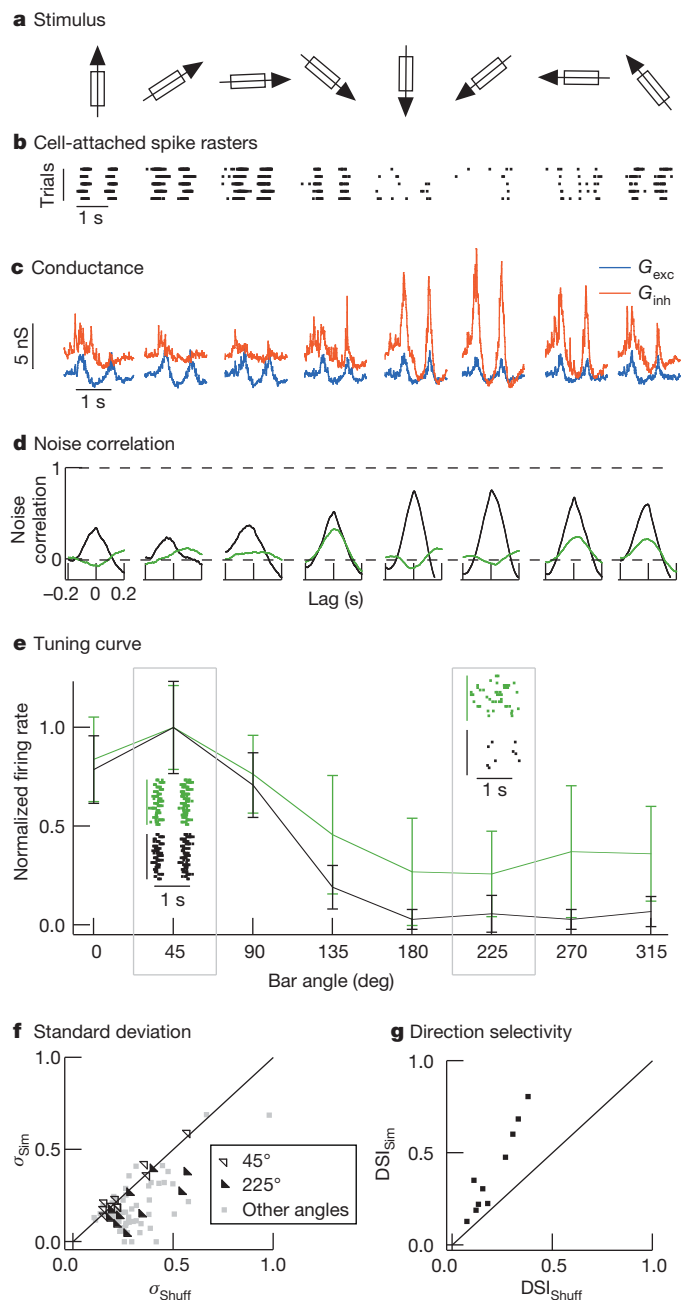
an increase in the excitatory conductance to be misinterpreted as an increase in both the excitatory and inhibitory conductances, thus leading to an artefactual correlation. A similar logic holds if a cell is held more negative than the reversal potential for inhibitory input. However, if anything the alternating-voltage technique fell short of the actual reversal potentials and hence underestimated the strength of noise correlations (Supplementary Fig. 3).

To determine the effect of covariation of excitatory and inhibitory synaptic inputs on a midretinal ganglion cell's response to physiological inputs, we compared the pattern of spikes produced by simultaneous (with noise correlations) and non-simultaneous (without noise correlations) conductances in dynamic-clamp experiments (Fig. 3c and Supplementary Fig. 4). The non-simultaneous conductances consisted of shuffled pairings of simultaneously recorded excitatory and inhibitory conductances; this procedure removed noise correlations while holding all other statistics constant. We compared the precision of the spike responses to the two sets of conductances by calculating the signal-to-noise ratio from repeated dynamic-clamp trials (Fig. 3d; see Methods). In all cases, the signal-to-noise ratio was higher for conductances with noise correlations (Fig. 3e). Quantifying the temporal precision of the spike responses using a spike distance metric<sup>22,23</sup> gave similar results (data not shown). Thus, the precision of a midretinal cell's output in response to light stimuli depends on the covariation of excitatory and inhibitory synaptic inputs.

Feed-forward synaptic inhibition can serve a more diverse functional role when the amplitude or timing of inhibitory input relative to excitatory input depends on the stimulus. For example, the ability of a subset of retinal ganglion cells to respond to the direction of a moving object<sup>24,25</sup> (Fig. 4a, b) relies on cancellation of excitatory input by inhibitory input in the non-preferred direction<sup>10</sup>. Covariation of excitatory and inhibitory synaptic inputs could make such a mechanism robust to noise, for example by preventing a larger-than-average excitatory synaptic event from overwhelming the corresponding inhibitory synaptic event and causing a response to movement in an inappropriate direction. To test this proposal, we recorded simultaneous conductances in mouse ON-OFF directionally selective ganglion cells (ON-OFF DSGCs) in response to a bar of light moving in different directions (Fig. 4a). Excitatory and inhibitory conductances showed strong noise correlations that were largely absent from non-simultaneous conductances (Fig. 4d; see Supplementary Fig. 5 for results from full-field light stimuli). Both excitatory and inhibitory conductances and the strength of the noise correlations depended on bar direction (Fig. 4c, d). For example, noise correlations in the non-preferred direction were three to four times stronger than those in the preferred direction. Furthermore, excitatory and inhibitory conductances showed near-perfect covariation in the non-preferred direction.

We tested the impact of noise correlations on direction tuning using simultaneous (with noise correlations) and non-simultaneous (without noise correlations) conductances in dynamic-clamp experiments; non-simultaneous conductances consisted of simultaneous conductances shuffled between trials but not bar directions. Both the mean and the standard deviation of the firing rate in the non-preferred direction were considerably higher for non-simultaneous conductances (Fig. 4e, f). The failure of a cell to attenuate its response reliably for movement in the non-preferred direction should negatively affect its ability to encode direction. Indeed, each recorded cell showed greater direction selectivity for the simultaneous conductances (Fig. 4g). Thus, the computation underlying directional selectivity depends on covariation of excitatory and inhibitory synaptic inputs and the resulting cancellation of noise shared between the circuits providing each type of input.

Computation in the retina follows a basic plan found in many other neural circuits: signals in a common population of inputs diverge to parallel and functionally dissimilar pathways, and integration of the signals from multiple parallel pathways governs the output of the circuit. Divergence into separate excitatory and inhibitory circuits is a prominent



**Figure 4 | Strength and impact of noise correlations in synaptic inputs to ON-OFF directionally selective ganglion cells.** **a**, A bar of light was moved in eight directions, at 45° increments in random order. **b**, Extracellular (cell-attached configuration) spike responses to the moving bar. **c**, Examples of simultaneously recorded conductances showing tuning of excitatory (blue) and inhibitory (red) conductances. **d**, Simultaneous conductances (black) show strong noise correlations that are largely absent from the non-simultaneous conductances (green). **e**, Normalized directional tuning (spike count versus direction) from a single dynamic-clamp experiment (mean  $\pm$  s.d.) for 20 trials of simultaneous or shuffled simultaneous conductances. Insets at 45° (preferred direction) and 225° (non-preferred direction) show spike rasters. **f**, Standard deviation of the normalized spike count is significantly smaller for simultaneous trials than for shuffled trials in non-preferred directions (135–315°;  $P < 0.05$ , 10 cells). Standard deviations in the preferred direction were similar. **g**, Direction selectivity index (DSI; see Methods) is  $2.0 \pm 0.2$  times larger for simultaneous conductances than for shuffled conductances (mean  $\pm$  s.e.m., 10 cells,  $P = 0.0002$ ).

example of such a motif. Noise in shared inputs naturally causes covariation of signals in the parallel pathways. The strength of such noise correlations will depend on cellular properties within the network<sup>15,16</sup>,



the stimulus delivered<sup>26</sup> (Fig. 4) and the state of the network<sup>27</sup>. Thus, excitatory and inhibitory inputs to cells in some, but not all, circuits are expected to show strong noise correlations, as indeed is the case in barrel cortex<sup>27,28</sup>. Here we put such noise correlations in the context of the coding of physiologically relevant stimuli. Our results reveal a critical role for noise correlations in maintaining appropriate cancellation of excitatory and inhibitory inputs and thus sharpening tuning to specific stimuli. This work provides an example of neurons that perform computations reliant on noise correlations. Given the prevalence of circuits in which feed-forward inhibition shapes neural responses<sup>9,11,12,19–21</sup>, noise correlations probably have a similar role in other neural circuits.

## METHODS SUMMARY

We took electrical recordings from midget ganglion cells in primate and ON–OFF DSGCs in mouse retinas using patch-clamp techniques as previously described<sup>23,29</sup>. Light stimuli were delivered from light-emitting diodes or an organic light-emitting diode monitor (eMagen). Mean light levels for all experiments were near 5,000 absorbed photons per cone per second.

The 10-ms cycle period during the simultaneous conductance recordings allows us to resolve input at 50 Hz and below. The fraction of the measured current variance at this cycle time was determined by calculating the fraction of the variance of the non-simultaneous (constant-voltage) conductances that can be accounted for by the variance of the simultaneous conductances.

Signal-to-noise ratios of spike outputs were calculated by forming spike trains of zeroes and ones from each trial, with 1-ms resolution. The mean and trial residuals of these spike trains were calculated and the power spectra of these functions were assessed and corrected for sample number bias<sup>30</sup>. Power spectra were integrated between 1 and 20 Hz and the result for the mean responses was divided by that for the residuals (Supplementary Fig. 6).

Spike number in ON–OFF DSGCs in response to the moving bar was summed over the entire duration of the bar's movement. The direction selectivity index<sup>10</sup> was calculated as  $DSI = |\sum \mathbf{v}_i / \sum \mathbf{r}_i|$ , where  $\mathbf{v}_i$  are vectors of lengths  $\mathbf{r}_i$ , equal to the normalized firing rate, and point in the direction of the moving bar that produced the presented conductances.

Current injected into a cell ( $I$ ) during dynamic-clamp experiments<sup>31</sup> was calculated as

$$I(t) = G_{\text{exc}}(t)(V(t - \Delta t) - E_{\text{exc}}) + G_{\text{inh}}(t)(V(t - \Delta t) - E_{\text{inh}})$$

where  $G_{\text{exc}}$  and  $G_{\text{inh}}$  are a pair of conductances recorded during light stimulation,  $V$  is the cell's membrane potential, and  $E_{\text{exc}}$  and  $E_{\text{inh}}$  are reversal potentials set respectively at 0 mV and –80 mV. Changing the inhibitory reversal potential,  $E_{\text{inh}}$ , to –50 mV did not substantially affect the results.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 27 July; accepted 11 October 2010.

Published online 5 December 2010.

- Shadlen, M. N. & Newsome, W. T. Noise, neural codes and cortical organization. *Curr. Opin. Neurobiol.* **4**, 569–579 (1994).
- Softky, W. R. & Koch, C. The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *J. Neurosci.* **13**, 334–350 (1993).
- Abbott, L. F. & Dayan, P. The effect of correlated variability on the accuracy of a population code. *Neural Comput.* **11**, 91–101 (1999).
- Romo, R., Hernandez, A., Zainos, A. & Salinas, E. Correlated neuronal discharges that increase coding efficiency during perceptual discrimination. *Neuron* **38**, 649–657 (2003).
- Salinas, E. & Sejnowski, T. J. Impact of correlated synaptic input on output firing rate and variability in simple neuronal models. *J. Neurosci.* **20**, 6193–6209 (2000).
- Dan, Y., Alonso, J. M., Usrey, W. M. & Reid, R. C. Coding of visual information by precisely correlated spikes in the lateral geniculate nucleus. *Nature Neurosci.* **1**, 501–507 (1998).

- Nirenberg, S., Carcieri, S. M., Jacobs, A. L. & Latham, P. E. Retinal ganglion cells act largely as independent encoders. *Nature* **411**, 698–701 (2001).
- Averbeck, B. B., Latham, P. E. & Pouget, A. Neural correlations, population coding and computation. *Nature Rev. Neurosci.* **7**, 358–366 (2006).
- Pouille, F. & Scanziani, M. Enforcement of temporal fidelity in pyramidal cells by somatic feed-forward inhibition. *Science* **293**, 1159–1163 (2001).
- Taylor, W. R. & Vaney, D. I. Diverse synaptic mechanisms generate direction selectivity in the rabbit retina. *J. Neurosci.* **22**, 7712–7720 (2002).
- Wilent, W. B. & Contreras, D. Dynamics of excitation and inhibition underlying stimulus selectivity in rat somatosensory cortex. *Nature Neurosci.* **8**, 1364–1370 (2005).
- Wehr, M. & Zador, A. M. Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature* **426**, 442–446 (2003).
- Leary, C. J., Edwards, C. J. & Rose, G. J. Midbrain auditory neurons integrate excitation and inhibition to generate duration selectivity: an *in vivo* whole-cell patch study in anurans. *J. Neurosci.* **28**, 5481–5493 (2008).
- Renart, A. *et al.* The asynchronous state in cortical circuits. *Science* **327**, 587–590 (2010).
- Trong, P. K. & Rieke, F. Origin of correlated activity between parasol retinal ganglion cells. *Nature Neurosci.* **11**, 1343–1351 (2008).
- de la Rocha, J., Doiron, B., Shea-Brown, E., Josic, K. & Reyes, A. Correlation between neural spike trains increases with firing rate. *Nature* **448**, 802–806 (2007).
- Dacey, D. M. & Petersen, M. R. Dendritic field size and morphology of midget and parasol ganglion cells of the human retina. *Proc. Natl Acad. Sci. USA* **89**, 9666–9670 (1992).
- Calkins, D. J. & Sterling, P. Absence of spectrally specific lateral inputs to midget ganglion cells in primate retina. *Nature* **381**, 613–615 (1996).
- Gabernet, L., Jadhav, S. P., Feldman, D. E., Carandini, M. & Scanziani, M. Somatosensory integration controlled by dynamic thalamocortical feed-forward inhibition. *Neuron* **48**, 315–327 (2005).
- Luna, V. M. & Schoppa, N. E. GABAergic circuits control input-spike coupling in the piriform cortex. *J. Neurosci.* **28**, 8851–8859 (2008).
- Mittmann, W., Koch, U. & Häusser, M. Feed-forward inhibition shapes the spike output of cerebellar Purkinje cells. *J. Physiol. (Lond.)* **563**, 369–378 (2005).
- Victor, J. D. & Purpura, K. P. Nature and precision of temporal coding in visual cortex: a metric-space analysis. *J. Neurophysiol.* **76**, 1310–1326 (1996).
- Murphy, G. J. & Rieke, F. Network variability limits stimulus-evoked spike timing precision in retinal ganglion cells. *Neuron* **52**, 511–524 (2006).
- Barlow, H. B., Hill, R. M. & Levick, W. R. Retinal ganglion cells responding selectively to direction and speed of image motion in the rabbit. *J. Physiol. (Lond.)* **173**, 377–407 (1964).
- Weng, S., Sun, W. & He, S. Identification of ON–OFF direction-selective ganglion cells in the mouse retina. *J. Physiol. (Lond.)* **562**, 915–923 (2005).
- Cohen, M. R. & Newsome, W. T. Context-dependent changes in functional circuitry in visual area MT. *Neuron* **60**, 162–173 (2008).
- Gentet, L. J., Avermann, M., Matyas, F., Staiger, J. F. & Petersen, C. C. Membrane potential dynamics of GABAergic neurons in the barrel cortex of behaving mice. *Neuron* **65**, 422–435 (2010).
- Okun, M. & Lampl, I. Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. *Nature Neurosci.* **11**, 535–537 (2008).
- Dunn, F. A., Lankheet, M. J. & Rieke, F. Light adaptation in cone vision involves switching between receptor and post-receptor sites. *Nature* **449**, 603–606 (2007).
- van Hateren, J. H. & Snippe, H. P. Information theoretical evaluation of parametric models of gain control in blowfly photoreceptor cells. *Vision Res.* **41**, 1851–1865 (2001).
- Sharp, A. A., O'Neil, M. B., Abbott, L. F. & Marder, E. Dynamic clamp: computer-generated conductances in real neurons. *J. Neurophysiol.* **69**, 992–995 (1993).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank D. Dacey, O. Packer, J. Crook, B. Peterson and T. Haun for providing primate tissue; P. Newman and E. Martinson for technical assistance; T. Azevedo, E. J. Chichilnisky, F. Dunn, G. Murphy, S. Kuo, E. Shea-Brown, M. Shadlen and W. Spain for comments on the manuscript and discussions. Support was provided by HHMI and NIH (EY-11850).

**Author Contributions** J.C. and F.R. designed and carried out the experiments, J.C. analysed the data and J.C. and F.R. wrote the paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to F.R. ([rieke@u.washington.edu](mailto:rieke@u.washington.edu)).

## METHODS

Electrical recordings were made from midget ganglion cells in primate and ON-OFF DSGCs in mouse retinas as previously described<sup>23,29</sup>. Midget ganglion cells were identified by their relatively sustained response to light steps and characteristic morphology<sup>17,29,32</sup>. ON-OFF DSGCs were identified by a combination of at least two of the following criteria: an on-off light response to a brief light step, a bistratified morphology and a directionally selective spike response.

Light stimuli were delivered from light-emitting diodes or an organic light-emitting diode monitor (eMagine). Mean light levels for all experiments were near 5000 absorbed photons per cone per second. Full-field stimuli consisted of 10 s of constant light followed by 10 s of 50%-contrast modulated light (low-pass-filtered at 60 Hz) repeated for 5–20 trials. Moving bars were 180  $\mu\text{m}$  wide, 720  $\mu\text{m}$  long, moved at 864  $\mu\text{m s}^{-1}$  along the long axis and had a contrast of between 100 and 150%.

For all recordings a flat-mounted piece of retina was superfused with warmed (31–34 °C) and oxygenated (5%  $\text{CO}_2$ , 95%  $\text{O}_2$ ) Ames solution. Midget cell dynamic-clamp experiments were performed with receptors mediating excitatory and inhibitory synaptic input blocked (10  $\mu\text{M}$  NBQX, 1  $\mu\text{M}$  strychnine, 10  $\mu\text{M}$  gabazine). Pipettes for voltage-clamp recordings were filled with a Cs-based internal solution (105 mM  $\text{CsCH}_3\text{SO}_3$ , 10 mM TEA-Cl, 20 mM HEPES, 10 mM EGTA, 5 mM Mg-ATP, 0.5 mM Tris-GTP and 2 mM QX-314, pH  $\sim 7.3$ ,  $\sim 280$  mosM). Pipettes for dynamic-clamp experiments were filled with a K-based internal solution (110 mM K aspartate, 1 mM MgCl, 10 mM HEPES, 5 mM NMDG, 0.5 mM  $\text{CaCl}_2$ , 10 mM phosphocreatine, 4 mM Mg-ATP and 0.5 mM Tris-GTP, pH  $\sim 7.2$ ,  $\sim 280$  mosM). Liquid junction potentials were  $\sim 10$  mV and were not compensated throughout the text. Low access resistance was critical, and only cells with access resistance below 20 M $\Omega$  were included for analysis. Access resistance was partially compensated for (75% for experiments using an Axopatch 200B amplifier; 50% compensation and prediction for experiments using a Multiclamp 700B amplifier). Conductances were derived from excitatory and inhibitory synaptic currents by dividing the currents by assumed driving forces corresponding to voltages of  $-62$  and  $+62$  mV, respectively.

Both ganglion cell types showed evidence for NMDA-receptor-mediated conductances (J-shaped  $I$ - $V$  plots that became linear in the presence of 10  $\mu\text{M}$  APV). The presence of an NMDA conductance could cause noise correlations to be substantially underestimated if the voltage is substantially below the excitatory reversal potential. However, we observed only a weak impact of this conductance

when noise correlations were compared before and after application of APV. Results from two cells recorded only in the presence of APV were included in the full data set.

The 10-ms cycle period during the simultaneous conductance recordings allows us to resolve input at 50 Hz and below. The fraction of the measured current variance at this cycle time was determined by calculating the fraction of the variance of the non-simultaneous (constant-voltage) conductances that can be accounted for by the variance of the simultaneous conductances.

Signal-to-noise ratios of spike outputs were calculated by forming spike trains of zeroes and ones from each trial, with 1-ms resolution. The mean and trial residuals of these spike trains were calculated and the power spectra of these functions were assessed and corrected for sample number bias<sup>30</sup>. Power spectra were integrated between 1 and 20 Hz and the result for the mean responses was divided by that for the residuals (Supplementary Fig. 6).

Spike number in ON-OFF DSGCs in response to the moving bar was summed over the entire duration of the bar's movement. The direction selectivity index<sup>10</sup> was calculated as  $\text{DSI} = |\sum \mathbf{v}_i / \sum \mathbf{r}_i|$ , where  $\mathbf{v}_i$  are vectors of lengths  $\mathbf{r}_i$ , equal to the normalized firing rate, and point in the direction of the moving bar that produced the presented conductances.

Current injected into a cell ( $I$ ) during dynamic-clamp experiments<sup>31</sup> was calculated as

$$I(t) = G_{\text{exc}}(t)(V(t - \Delta t) - E_{\text{exc}}) + G_{\text{inh}}(t)(V(t - \Delta t) - E_{\text{inh}})$$

where  $G_{\text{exc}}$  and  $G_{\text{inh}}$  are a pair of conductances recorded during light stimulation,  $V$  is the cell's membrane potential, and  $E_{\text{exc}}$  and  $E_{\text{inh}}$  are reversal potentials set respectively at 0 mV and  $-80$  mV. Changing the inhibitory reversal potential,  $E_{\text{inh}}$ , to  $-50$  mV did not substantially affect the results.

Correlations were calculated using the 'xcov' function in MATLAB, release 2009a (MathWorks) and normalized using the 'coef' option. Briefly, this function calculates the cross-correlation after subtracting the means from each trial and normalizes by the geometric mean of the autocorrelation (see Supplementary Information, equation (2.1)).

32. Polyak, S. & Willmer, E. N. Retinal structure and colour vision. *Doc. Ophthalmol.* **3**, 24–56 (1949).

# COT drives resistance to RAF inhibition through MAP kinase pathway reactivation

Cory M. Johannessen<sup>1,2\*</sup>, Jesse S. Boehm<sup>1\*</sup>, So Young Kim<sup>1,2,3†</sup>, Sapana R. Thomas<sup>1,2</sup>, Leslie Wardwell<sup>2</sup>, Laura A. Johnson<sup>1,2</sup>, Caroline M. Emery<sup>2</sup>, Nicolas Stransky<sup>1</sup>, Alexandria P. Cogdill<sup>4</sup>, Jordi Barretina<sup>1,2,5</sup>, Giordano Caponigro<sup>6</sup>, Haley Hieronymus<sup>1,7,8</sup>, Ryan R. Murray<sup>3,9,10</sup>, Kourosh Salehi-Ashtiani<sup>3,9,10</sup>, David E. Hill<sup>3,9,10</sup>, Marc Vidal<sup>3,9,10</sup>, Jean J. Zhao<sup>9,11</sup>, Xiaoping Yang<sup>1</sup>, Ozan Alkan<sup>1</sup>, Sungjoon Kim<sup>12</sup>, Jennifer L. Harris<sup>12</sup>, Christopher J. Wilson<sup>6</sup>, Vic E. Myer<sup>6</sup>, Peter M. Finan<sup>6</sup>, David E. Root<sup>1</sup>, Thomas M. Roberts<sup>9</sup>, Todd Golub<sup>1,5,8</sup>, Keith T. Flaherty<sup>4</sup>, Reinhard Dummer<sup>13</sup>, Barbara L. Weber<sup>6</sup>, William R. Sellers<sup>6</sup>, Robert Schlegel<sup>6</sup>, Jennifer A. Wargo<sup>4</sup>, William C. Hahn<sup>1,2,3,5</sup> & Levi A. Garraway<sup>1,2,5</sup>

**Oncogenic mutations in the serine/threonine kinase B-RAF (also known as BRAF) are found in 50–70% of malignant melanomas<sup>1</sup>. Pre-clinical studies have demonstrated that the B-RAF(V600E) mutation predicts a dependency on the mitogen-activated protein kinase (MAPK) signalling cascade in melanoma<sup>2–6</sup>—an observation that has been validated by the success of RAF and MEK inhibitors in clinical trials<sup>7–9</sup>. However, clinical responses to targeted anticancer therapeutics are frequently confounded by *de novo* or acquired resistance<sup>10–12</sup>. Identification of resistance mechanisms in a manner that elucidates alternative ‘druggable’ targets may inform effective long-term treatment strategies<sup>13</sup>. Here we expressed ~600 kinase and kinase-related open reading frames (ORFs) in parallel to interrogate resistance to a selective RAF kinase inhibitor. We identified *MAP3K8* (the gene encoding COT/Tpl2) as a MAPK pathway agonist that drives resistance to RAF inhibition in B-RAF(V600E) cell lines. COT activates ERK primarily through MEK-dependent mechanisms that do not require RAF signalling. Moreover, COT expression is associated with *de novo* resistance in B-RAF(V600E) cultured cell lines and acquired resistance in melanoma cells and tissue obtained from relapsing patients following treatment with MEK or RAF inhibitors. We further identify combinatorial MAPK pathway inhibition or targeting of COT kinase activity as possible therapeutic strategies for reducing MAPK pathway activation in this setting. Together, these results provide new insights into resistance mechanisms involving the MAPK pathway and articulate an integrative approach through which high-throughput functional screens may inform the development of novel therapeutic strategies.**

To identify kinases capable of circumventing RAF inhibition, we assembled and stably expressed 597 sequence-validated kinase ORF clones representing ~75% of annotated kinases (Center for Cancer Systems Biology (CCSB)/Broad Institute Kinase ORF Collection) in A375, a B-RAF(V600E) malignant melanoma cell line that is sensitive to the RAF kinase inhibitor PLX4720<sup>14</sup> (Fig. 1a, b, Supplementary Table 1 and Supplementary Fig. 2). ORF-expressing cells treated with 1  $\mu$ M PLX4720 were screened for viability relative to untreated cells and normalized to an assay-specific positive control, MEK1(S218/222D) (MEK1<sup>DD</sup>)<sup>15</sup> (Supplementary Table 2 and summarized in Supplementary Fig. 1). Nine ORFs conferred resistance at levels exceeding two standard deviations from the mean (Fig. 1b and Supplementary

Table 2) and were selected for follow-up analysis (Supplementary Fig. 3). Three of the nine candidate ORFs were receptor tyrosine kinases, underscoring the potential of this class of kinases to engage resistance pathways. Resistance effects were validated and prioritized across a multi-point PLX4720 drug concentration scale in the B-RAF(V600E) cell lines A375 and SKMEL28. The Ser/Thr MAP kinase kinase kinases (MAP3Ks) *MAP3K8* (COT/Tpl2) and *RAF1* (C-RAF) emerged as top candidates from both cell lines; these ORFs shifted the PLX4720 half-maximal growth inhibitory concentration (GI<sub>50</sub>) by 10–600-fold without affecting viability (Supplementary Table 3 and Supplementary Figs 4 and 5). Both COT and C-RAF reduced sensitivity to PLX4720 in multiple B-RAF(V600E) cell lines (Fig. 1c) confirming the ability of these kinases to mediate resistance to RAF inhibition.

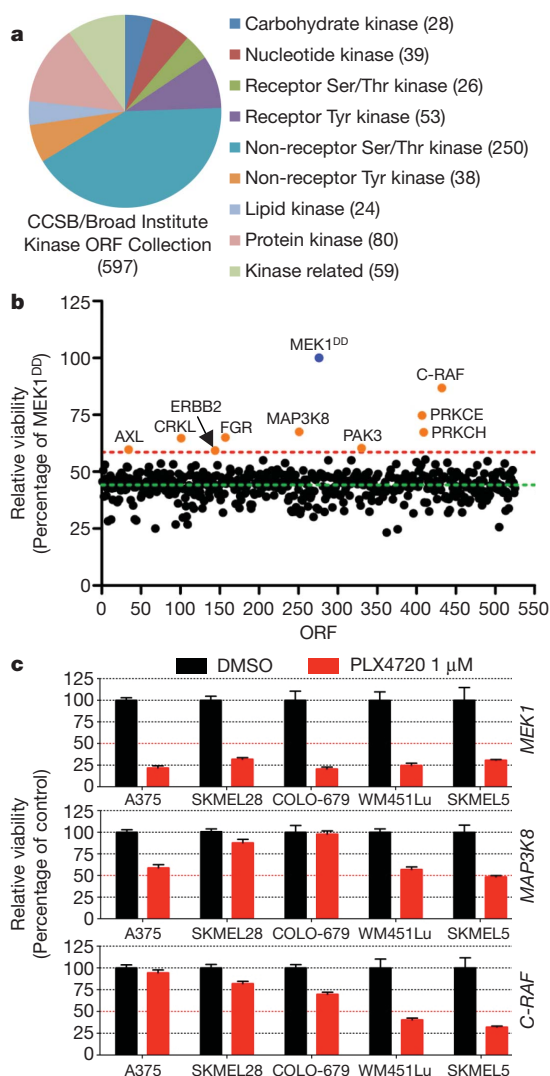
Next, we tested whether overexpression of these genes was sufficient to activate the MAPK pathway. At baseline, COT expression increased ERK phosphorylation in a manner comparable to MEK1<sup>DD</sup>, consistent with MAP kinase pathway activation (Fig. 2a and Supplementary Fig. 6). Overexpression of wild-type COT or C-RAF resulted in constitutive phosphorylation of ERK and MEK in the presence of PLX4720, whereas kinase-dead derivatives had no effect (Fig. 2a and Supplementary Fig. 7). Based on these results, we proposed that COT and C-RAF drive resistance to RAF inhibition predominantly through re-activation of MAPK signalling. Notably, of the nine candidate ORFs from our initial screen, a subset (three) did not show persistent ERK/MEK phosphorylation following RAF inhibition, suggesting MAPK pathway-independent alteration of drug sensitivity (Supplementary Fig. 8).

Several groups have shown that C-RAF activation and heterodimerization with B-RAF constitute critical components of the cellular response to B-RAF inhibition<sup>16–19</sup>. In A375 cells, endogenous C-RAF–B-RAF heterodimers were measurable and inducible following treatment with PLX4720 (Supplementary Fig. 9). However, endogenous C-RAF phosphorylation at S338—an event required for C-RAF activation—remained low (Supplementary Fig. 9). In contrast, ectopically expressed C-RAF was phosphorylated on S338 (Supplementary Fig. 9) and its PLX4720 resistance phenotype was associated with sustained MEK/ERK activation (Fig. 2a, Supplementary Fig. 9). Moreover, ectopic expression of a high-activity C-RAF truncation mutant (C-RAF(W22)) was more effective than wild-type C-RAF in mediating PLX4720 resistance and ERK activation (Supplementary Fig. 10), further indicating that

<sup>1</sup>Broad Institute of Harvard and Massachusetts Institute of Technology, 7 Cambridge Center, Cambridge, Massachusetts 02142, USA. <sup>2</sup>Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, 44 Binney Street, Boston, Massachusetts 02115, USA. <sup>3</sup>Center for Cancer Systems Biology (CCSB), Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA. <sup>4</sup>Division of Surgical Oncology, Medical Oncology and Dermatology, Massachusetts General Hospital, 55 Fruit Street, Boston, Massachusetts 02114, USA. <sup>5</sup>Center for Cancer Genome Discovery, Dana-Farber Cancer Institute, Harvard Medical School, 44 Binney Street, Boston, Massachusetts 02115, USA. <sup>6</sup>Novartis Institutes for Biomedical Research, 250 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA. <sup>7</sup>Current address: Human Oncology and Pathogenesis Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. <sup>8</sup>Department of Pediatric Oncology, Dana-Farber Cancer Institute, Harvard Medical School, 44 Binney Street, Boston, Massachusetts 02115, USA. <sup>9</sup>Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA. <sup>10</sup>Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>11</sup>Department of Pathology, Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>12</sup>Genomics Institute of the Novartis Research Foundation, San Diego, California 92121, USA. <sup>13</sup>Department of Dermatology, University Hospital of Zurich, Zurich CH-8091, Switzerland. <sup>†</sup>Present address: Duke Institute for Genome Sciences and Policy, Duke University Medical Center, Durham, North Carolina 27710, USA (S.Y.K.).

\*These authors contributed equally to this work.





**Figure 1 | An ORF-based functional screen identifies COT and C-RAF kinases as drivers of resistance to B-RAF inhibition** **a**, Overview of the CCSB/Broad Institute Kinase ORF collection. Kinase classification and number of kinases per classification are noted. **b**, A375 cells expressing the CCSB/Broad Institute Kinase ORF collection were assayed for relative viability in 1  $\mu$ M PLX4720 and normalized to constitutively active MEK1 (MEK1<sup>DD</sup>). Nine ORFs (orange disks) scored 2 standard deviations (red dashed line, 58.64%) from the mean of all ORFs (green dashed line, 44.26%). **c**, Indicated ORFs were expressed in five B-RAF(V600E) cell lines and treated with DMSO or 1  $\mu$ M PLX4720. Viability (relative to DMSO) was quantified after 4 days. Error bars represent standard deviation between replicates ( $n = 6$ ).

elevated C-RAF activity may direct resistance to this agent. Consistent with this model, oncogenic alleles of *NRAS* and *KRAS* conferred PLX4720 resistance in A375 cells (Fig. 2b) and yielded sustained C-RAF(S338) and ERK phosphorylation in the context of drug treatment (Fig. 2c). Thus, although genetic alterations that engender C-RAF activation (for example, oncogenic RAS mutations) tend to show mutual exclusivity with B-RAF(V600E) mutation, such co-occurring events<sup>20,21</sup> might be favoured in the context of acquired resistance to B-RAF inhibition.

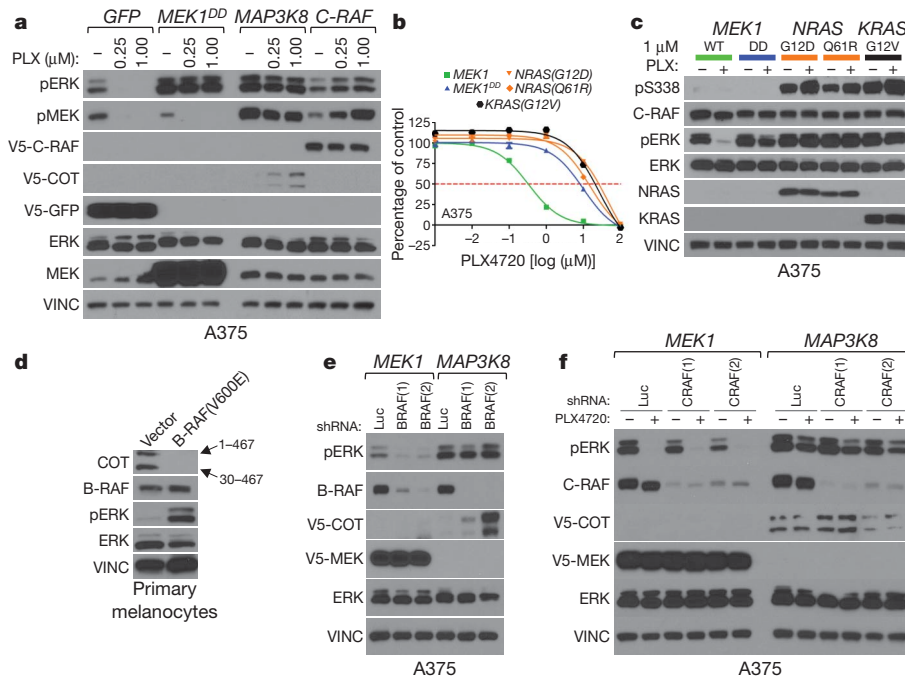
To investigate the role of COT in melanoma, we first determined its expression in human melanocytes. We found that primary immortalized melanocytes (B-RAF wild-type) expressed COT (Fig. 2d), although ectopic B-RAF(V600E) expression reduced *MAP3K8* mRNA levels (Supplementary Fig. 11) and rendered COT protein undetectable (Fig. 2d). Conversely, whereas ectopically expressed COT was only weakly detectable in A375 cells (Fig. 2a, e), short hairpin RNA (shRNA)-mediated depletion of endogenous B-RAF(V600E) caused an increase in COT

protein levels that correlated with the extent of B-RAF knockdown (Fig. 2e). Moreover, treatment of COT-expressing A375 cells with PLX4720 led to a dose-dependent increase in COT protein (Fig. 2a) without affecting ectopic *MAP3K8* mRNA levels (Supplementary Fig. 11). Thus, oncogenic B-RAF may antagonize COT expression largely through altered protein stability (Fig. 2a, d, e and Supplementary Fig. 11), and B-RAF inhibition may potentiate the outgrowth of COT-expressing cells during the course of treatment. Notably, neither C-RAF nor B-RAF alone or in combination was required for ERK phosphorylation in the context of COT expression, even in the presence of PLX4720 (Fig. 2e, f and Supplementary Fig. 12), suggesting that COT expression is sufficient to induce MAP kinase pathway activation in a RAF-independent manner.

We predicted that cell lines expressing elevated COT in a B-RAF(V600E) background should show *de novo* resistance to PLX4720 treatment. To identify such instances, we screened a panel of cell lines for evidence of *MAP3K8* copy number gains coincident with the B-RAF(V600E) mutation. Of 534 cell lines that had undergone copy number analysis and mutation profiling, 38 cell lines (7.1%) contained the B-RAF(V600E) mutation. Within this subgroup, two cell lines—OUMS-23 (colon cancer) and RPMI-7951 (melanoma)—also showed evidence of chromosomal copy gains spanning the *MAP3K8* locus (Fig. 3a and Supplementary Fig. 13) and robust COT protein expression (Fig. 3b and Supplementary Fig. 14). We also screened a panel of melanoma short-term cultures for COT protein expression. Only one of these lines expressed COT: M307, a short-term culture derived from a B-RAF(V600E) tumour that developed resistance to allosteric MEK inhibition following initial disease stabilization<sup>15</sup> (Fig. 3c). All three cell lines were refractory to PLX4720 treatment, with GI<sub>50</sub> values in the range of 8–10  $\mu$ M (Fig. 3d), and showed sustained ERK phosphorylation in the context of B-RAF inhibition (Fig. 3e, f). OUMS-23 and RPMI-7951 are MAPK pathway inhibitor-naïve cell lines, implying that COT may confer *de novo* resistance to RAF inhibition (a phenomenon observed in ~10% of B-RAF(V600E) melanomas<sup>7</sup>).

Next, we examined COT expression in the context of resistance to the clinical RAF inhibitor PLX4032 by obtaining biopsy material from three patients with metastatic, B-RAF(V600E) melanoma. Each case consisted of frozen, lesion-matched biopsy material obtained before and during treatment ('pre-treatment' and 'on-treatment'; Fig. 3g and Supplementary Table 4); additionally, one sample contained two independent biopsy specimens from the same relapsing tumour site ('post-relapse'; Fig. 3g). Consistent with the experimental models presented above, quantitative real-time PCR with reverse transcription (qRT-PCR) analysis revealed increased *MAP3K8* mRNA expression concurrent with PLX4032 treatment in two of three cases. *MAP3K8* mRNA levels were further increased in a relapsing specimen relative to its pre-treatment and on-treatment counterparts (Fig. 3g, Patient 1). An additional, unmatched relapsed malignant melanoma biopsy showed elevated *MAP3K8* mRNA expression comparable to levels observed in RAF inhibitor-resistant, *MAP3K8*-amplified cell lines (Supplementary Fig. 15). This specimen also exhibited robust MAPK pathway activation and elevated expression of B-RAF, C-RAF and COT relative to matched normal skin or B-RAF(V600E) cell lines (Supplementary Fig. 15). Sequencing studies of this tumour revealed no additional mutations in B-RAF, *NRAS* or *KRAS* (data not shown). These analyses provided clinical evidence that COT-dependent mechanisms may be operant in at least some PLX4032-resistant malignant melanomas.

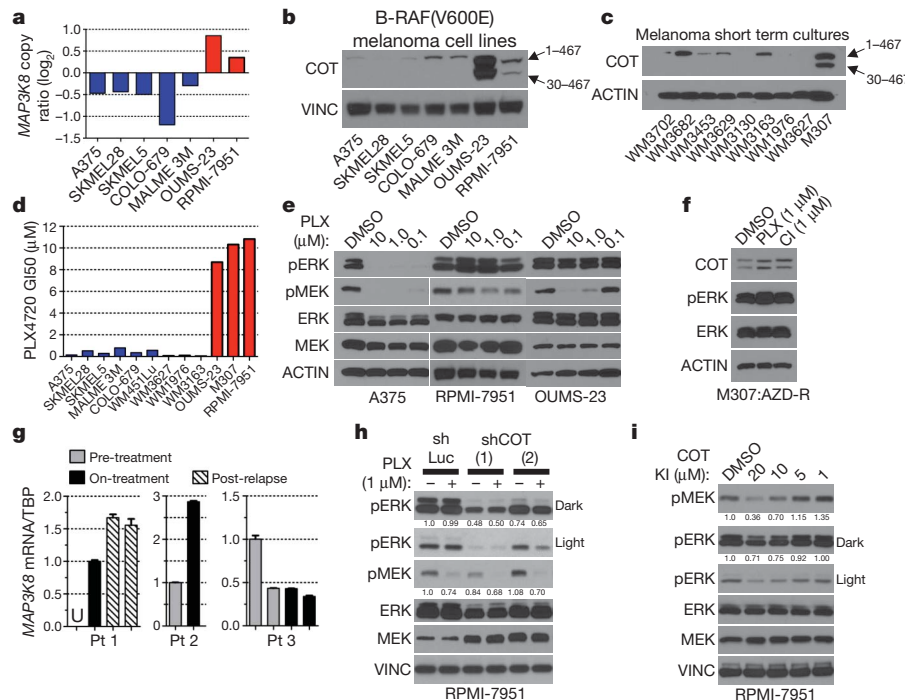
To determine if COT might actively regulate MEK/ERK phosphorylation in B-RAF(V600E) cells that harbour naturally elevated COT expression, we introduced shRNA constructs targeting *MAP3K8*/COT into RPMI-7951 cells. Depletion of COT suppressed RPMI-7951 viability (Supplementary Fig. 16) and decreased ERK phosphorylation (Fig. 3h), implying that targeting COT kinase activity might suppress MEK/ERK phosphorylation in cancer cells with COT overexpression or amplification. Treatment of RPMI-7951 cells with a small molecule COT kinase inhibitor<sup>22–24</sup> resulted in dose-dependent suppression of



**Figure 2 | Resistance to B-RAF inhibition via MAPK pathway activation**

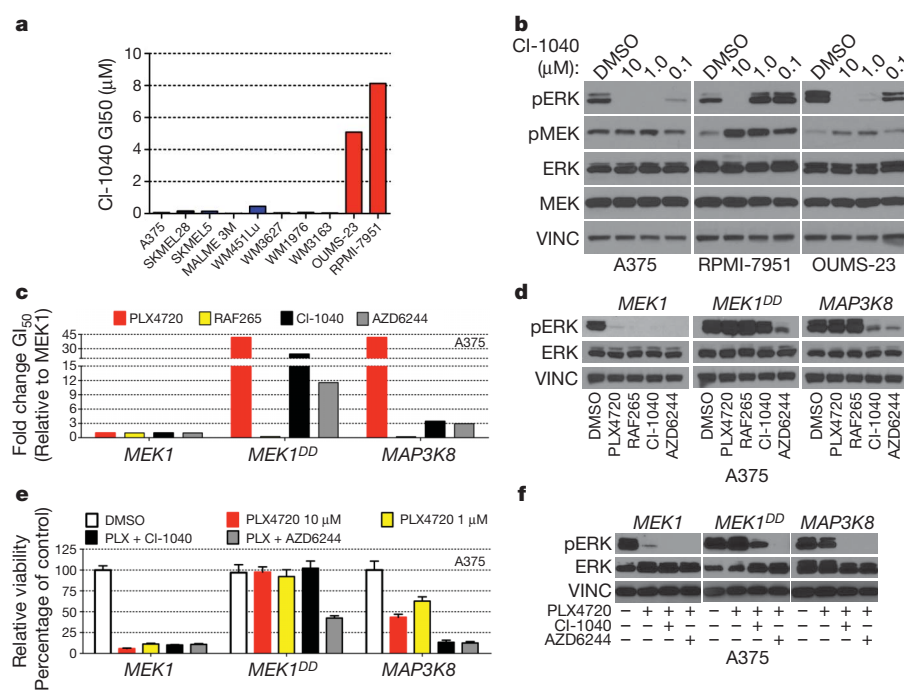
**a**, Indicated ORFs were expressed in A375. Levels of phosphorylated MEK and ERK were assayed after 18 h treatment with DMSO (–) or PLX4720 (concentration noted). GFP, green fluorescent protein; V5-C-RAF, V5-COT, V5-GFP, V5 epitope-tagged C-RAF, COT and GFP, respectively; PLX, PLX4720. **b**, Proliferation of A375 expressing indicated ORFs. Error bars represent standard deviation between replicates ( $n = 6$ ). **c**, C-RAF (S338) and ERK phosphorylation in lysates from A375 expressing indicated ORFs. Vinc, vinculin; WT, wild type; pS338, C-RAF phosphorylated on Ser338. **d**, COT

expression in lysates from immortalized primary melanocytes expressing BRAF(V600E) or empty vector. MAP3K8 mRNA has an internal start codon (30<sup>M</sup>) resulting in two protein products of different lengths; amino acids 1–467 or 30–467, noted with arrows. **e**, COT expression and ERK phosphorylation in lysates from A375 expressing indicated ORFs following shRNA-mediated B-RAF depletion (shBRAF) relative to control shRNA (shLuc). **f**, ERK phosphorylation in lysates from A375 expressing indicated ORFs following shRNA-mediated C-RAF depletion (shCRAF) or control shRNA (shLuc), after 18 h treatment with DMSO (–) or 1 μM PLX4720 (+).



**Figure 3 | COT expression predicts resistance to B-RAF inhibition in cancer cell lines** **a**, MAP3K8 copy numbers. Red bars, MAP3K8 amplification; blue bars, non-amplified COT. **b**, COT expression in B-RAF(V600E) cell lines and **c**, short-term cultures. **d**, PLX4720 GI<sub>50</sub> in B-RAF(V600E) cell lines. Colours as in **a**. **e**, MEK and ERK phosphorylation after treatment with DMSO or PLX4720 (concentration indicated). **f**, ERK phosphorylation in M307 lysates (AZD-R; AZD6244-resistant) treated with DMSO or 1 μM PLX4720 (PLX) or CI-1040 (CI). **g**, MAP3K8 mRNA expression (qRT-PCR) in patient/lesion-matched

PLX4032-treated metastatic melanoma tissue samples. Patients 1 and 3 had multiple biopsies from the same lesion. Error bars represent s.e.m. ( $n = 3$ ). Pt, patient; U, undetermined/undetectable; TBP, TATA binding protein. **h**, ERK and MEK phosphorylation in RPMI-7951 following shRNA-mediated COT depletion (shCOT) versus control (shLuc) and treatment with DMSO (–) or 1 μM PLX4720 (+). ERK and MEK phosphorylation are quantified. **i**, ERK and MEK phosphorylation in RPMI-7951 after 1 h treatment with a small molecule COT kinase inhibitor. ERK and MEK phosphorylation are quantified.



**Figure 4 | COT-expressing B-RAF(V600E) cell lines exhibit resistance to allosteric MEK inhibitors** **a**, CI-1040 GI<sub>50</sub> in a panel of B-RAF(V600E) cell lines. Red bars, COT expression/amplification; blue bars, undetectable/non-amplified COT. **b**, MEK and ERK phosphorylation in lysates from indicated cell lines treated with DMSO or CI-1040 (concentration noted). **c**, Fold change (relative to MEK1) GI<sub>50</sub> of A375 ectopically expressing the indicated ORFs for PLX4720, RAF265, CI-1040 and AZD6244. **d**, ERK phosphorylation in A375

MEK and ERK phosphorylation, providing additional evidence that COT contributes to MEK/ERK activation in these cells (Fig. 3i).

We then considered whether COT-expressing cancer cells remain sensitive to MAPK pathway inhibition at a target downstream of COT or RAF. Here, we queried the OUMS-23 and RPMI-7951 cell lines for sensitivity to the MEK1/2 inhibitor CI-1040. Interestingly, both cell lines were refractory to MEK inhibition (Fig. 4a) and displayed sustained ERK phosphorylation even at 1 μM CI-1040 (Fig. 4b). Ectopic COT expression in A375 and SKMEL28 cells also conferred decreased sensitivity to the MEK inhibitors CI-1040 and AZD6244, suggesting that COT expression alone was sufficient to induce this phenotype (Fig. 4c, d and Supplementary Fig. 17). Similar to results observed with pharmacological MEK inhibitors, MEK1/2 knockdown only modestly suppressed COT-mediated ERK phosphorylation in A375 cells (Supplementary Fig. 18). In accordance with prior observations<sup>25</sup>, these data raised the possibility that COT may activate ERK through MEK-independent as well as MEK-dependent mechanisms. To test this hypothesis directly, we performed an *in vitro* kinase assay using recombinant COT and ERK1. Indeed, recombinant COT induced pThr 202/Tyr 204 phosphorylation of ERK1 *in vitro* (Supplementary Fig. 18), indicating that in certain contexts COT expression may potentiate ERK activation in a MEK-independent manner.

In experimental models, the use of RAF and MEK inhibitors in combination can override resistance to single agents<sup>15</sup>. We therefore reasoned that combined RAF/MEK inhibition might circumvent COT-driven resistance. In the setting of ectopic COT expression, exposure to AZD6244 or CI-1040 in combination with PLX470 (1 μM each) reduced cell growth and pERK expression more effectively than did single-agent PLX4720, even at concentrations of 10 μM (Fig. 4e, f and Supplementary Fig. 19). These data underscore the importance of this pathway in B-RAF(V600E) tumour cells and support earlier findings<sup>15</sup> that dual B-RAF/MEK inhibition may help circumvent resistance to RAF inhibitors.

B-RAF mutations are found in ~8% of all cancers and at high frequencies in malignant melanoma, colon and thyroid cancers<sup>1</sup>.

expressing indicated ORFs following treatment with DMSO or 1 μM of PLX4720, RAF265, CI-1040 or AZD6244. **e**, Viability of A375 expressing the indicated ORFs and treated with DMSO, PLX4720 (concentration indicated) and PLX4720 in combination with CI-1040 or AZD6244 (all 1 μM). Error bars represent the standard deviation ( $n = 6$ ). **f**, ERK phosphorylation in A375 expressing indicated ORFs following treatment with DMSO, PLX4720 (1 μM) or PLX4720 in combination with CI-1040 or AZD6244 (all 1 μM).

The clinical promise of selective RAF inhibitors has widespread ramifications for patient treatment, yet single-agent targeted therapy is almost invariably followed by relapse due to acquired drug resistance. Our results suggest that ORF-based, systematic functional screening may offer a powerful means to identify clinically relevant resistance mechanisms that also specify novel treatment strategies. In particular, resistance to RAF inhibition can be achieved by multiple MAP3K-dependent mechanisms of MEK/ERK reactivation but might be intercepted through combined therapeutic modalities for MAPK pathway inhibition (for example, RAF/MEK or RAF/COT combinations). Future systematic drug resistance studies may be expanded to a genome scale that encompasses many compounds, thereby enabling comprehensive identification of both therapy-specific resistance genes and drug targets of novel therapeutics.

## METHODS SUMMARY

The arrayed, lentiviral ORF screen was performed as described previously<sup>26</sup>. Effects of individual ORFs on drug resistance were determined by measuring differential viability (ratio of raw viability in 1 μM PLX4720 over control) and subsequent normalization to an assay-specific positive control, MEK1<sup>DD</sup>. Secondary screens were performed with the top nine candidate ORFs in 96-well format in A375 and SKMEL28 cells. Prioritization was accomplished via generation of a GI<sub>50</sub> for each ORF across a multi-point PLX4720 concentration range in both cell lines. The effects of identified resistance ORFs on MAPK pathway activation were demonstrated using both biochemical and cell biological approaches. Cell line copy number data was obtained as previously described<sup>27</sup>. Detailed descriptions of all procedures are included in Methods.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 4 March; accepted 25 October 2010.

Published online 24 November 2010.

1. Davies, H. *et al.* Mutations of the *BRAF* gene in human cancer. *Nature* **417**, 949–954 (2002).



2. Hoefflich, K. P. *et al.* Antitumor efficacy of the novel RAF inhibitor GDC-0879 is predicted by BRAF<sup>V600E</sup> mutational status and sustained extracellular signal-regulated kinase/mitogen-activated protein kinase pathway suppression. *Cancer Res.* **69**, 3042–3051 (2009).
3. McDermott, U. *et al.* Identification of genotype-correlated sensitivity to selective kinase inhibitors by using high-throughput tumor cell line profiling. *Proc. Natl Acad. Sci. USA* **104**, 19936–19941 (2007).
4. Solit, D. B. *et al.* BRAF mutation predicts sensitivity to MEK inhibition. *Nature* **439**, 358–362 (2006).
5. Wan, P. T. *et al.* Mechanism of activation of the RAF-ERK signaling pathway by oncogenic mutations of B-RAF. *Cell* **116**, 855–867 (2004).
6. Wellbrock, C. *et al.* V599E-B-RAF is an oncogene in melanocytes. *Cancer Res.* **64**, 2338–2342 (2004).
7. Flaherty, K. T. *et al.* Inhibition of mutated, activated BRAF in metastatic melanoma. *N. Engl. J. Med.* **363**, 809–819 (2010).
8. Infante, J. R. *et al.* Safety and efficacy results from the first-in-human study of the oral MEK 1/2 inhibitor GSK1120212. *J. Clin. Oncol.* **28** (suppl.), 2503 (2010).
9. Schwartz, G. K. *et al.* A phase I study of XL281, a selective oral RAF kinase inhibitor, in patients (Pts) with advanced solid tumors. *J. Clin. Oncol.* **27** (suppl.), 3513 (2009).
10. Engelman, J. A. *et al.* MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* **316**, 1039–1043 (2007).
11. Gorre, M. E. *et al.* Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. *Science* **293**, 876–880 (2001).
12. Heinrich, M. C. *et al.* Molecular correlates of imatinib resistance in gastrointestinal stromal tumors. *J. Clin. Oncol.* **24**, 4764–4774 (2006).
13. Daub, H., Specht, K. & Ullrich, A. Strategies to overcome resistance to targeted protein kinase inhibitors. *Nature Rev. Drug Discov.* **3**, 1001–1010 (2004).
14. Tsai, J. *et al.* Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity. *Proc. Natl Acad. Sci. USA* **105**, 3041–3046 (2008).
15. Emery, C. M. *et al.* MEK1 mutations confer resistance to MEK and B-RAF inhibition. *Proc. Natl Acad. Sci. USA* **106**, 20411–20416 (2009).
16. Hatzivassiliou, G. *et al.* RAF inhibitors prime wild-type RAF to activate the MAPK pathway and enhance growth. *Nature* **464**, 431–435 (2010).
17. Heidorn, S. J. *et al.* Kinase-dead BRAF and oncogenic RAS cooperate to drive tumor progression through CRAF. *Cell* **140**, 209–221 (2010).
18. Karreth, F. A., DeNicola, G. M., Winter, S. P. & Tuveson, D. A. C-Raf inhibits MAPK activation and transformation by B-Raf<sup>V600E</sup>. *Mol. Cell* **36**, 477–486 (2009).
19. Poulikakos, P. I., Zhang, C., Bollag, G., Shokat, K. M. & Rosen, N. RAF inhibitors transactivate RAF dimers and ERK signalling in cells with wild-type BRAF. *Nature* **464**, 427–430 (2010).
20. Edlundh-Rose, E. *et al.* NRAS and BRAF mutations in melanoma tumours in relation to clinical characteristics: a study based on mutation screening by pyrosequencing. *Melanoma Res.* **16**, 471–478 (2006).
21. Seth, R. *et al.* Concomitant mutations and splice variants in KRAS and BRAF demonstrate complex perturbation of the Ras/Raf signalling pathway in advanced colorectal cancer. *Gut* **58**, 1234–1241 (2009).
22. George, D. *et al.* Discovery of thieno[2,3-c]pyridines as potent COT inhibitors. *Bioorg. Med. Chem. Lett.* **18**, 4952–4955 (2008).
23. Hirata, K. *et al.* Inhibition of tumor progression locus 2 protein kinase suppresses receptor activator of nuclear factor- $\kappa$ B ligand-induced osteoclastogenesis through down-regulation of the c-Fos and nuclear factor of activated T cells c1 genes. *Biol. Pharm. Bull.* **33**, 133–137 (2010).
24. Lee, K. M., Lee, K. W., Bode, A. M., Lee, H. J. & Dong, Z. Tpl2 is a key mediator of arsenite-induced signal transduction. *Cancer Res.* **69**, 8043–8049 (2009).
25. Tsatsanis, C., Patriotis, C. & Tschlis, P. N. Tpl-2 induces IL-2 expression in T-cell lines by triggering multiple signaling pathways that activate NFAT and NF- $\kappa$ B. *Oncogene* **17**, 2609–2618 (1998).
26. Barbie, D. A. *et al.* Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* **462**, 108–112 (2009).
27. Beroukhi, R. *et al.* The landscape of somatic copy-number alteration across human cancers. *Nature* **463**, 899–905 (2010).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank members of the Broad Institute/Novartis Cancer Cell Line Encyclopedia (CCLE) for contributing cell line genomic data, expression data and pharmacological cell line sensitivity data; J. Thibault and A. Shipway for CCLE related tissue culture; R. Depinho, G. Dunn, S. Ethier, H. Greulich, A. Henderson, D. Kaplan, R. Levine, C. Miller, H. Piwnica-Worms, H. Suzuki, M. Vigny, D. Vollrath and the Harvard Institute of Proteomics for contributing templates for the kinase collection; J. Du and D. B. Wheeler for assistance with functional testing of kinases; D. A. Barbie for helpful discussions, S. E. Moody and H. W. Cheung for technical assistance, and J. K. Grenier and S. J. Silver for compiling and annotating the list of kinases. This work was supported by the NIH Director's New Innovator Award (L.A.G.), grants from the Novartis Institutes for Biomedical Research (L.A.G.), Melanoma Research Alliance (L.A.G.), Starr Cancer Consortium (L.A.G.) the US National Cancer Institute (R33 CA128625, RC2 CA148268) (W.C.H.), NIH (CA134502) (J.J.Z.) the Swiss National Foundation (310040-103671) (R.D.), the Gottfried and Julia Bangerter Rhyner Stiftung (R.D.) the Ellison Foundation (M.V. and CCSB) and institute sponsored research funds from the DFCI Strategic Initiative (M.V. and CCSB).

**Author Contributions** C.M.J., J.S.B. and L.A.G. designed the experiments, with help from C.M.E. C.M.J., S.Y.K. and L.W. performed the primary screen, supervised by W.C.H. L.A.J. helped perform drug sensitivity profiling. J.S.B. designed and created the CCSB/Broad Institute Kinase ORF Collection in collaboration with S.R.T., H.H., R.R.M., K.S.-A., J.J.Z., M.V., T.M.R., T.G., D.E.H. and W.C.H. Additional help with ORF experiments was provided by X.Y., D.E.R. and O.A. Clinical samples were collected or experiments performed by C.M.J., A.P.C., K.T.F., R.D. and J.A.W. Large scale cell genomic and expression profiling along with pharmacological screening efforts were designed and data analysed by N.S., J.B., G.C., S.K., J.H., C.J.W., V.E.M., P.M.F., B.L.W., W.R.S., R.S. and L.A.G. C.M.J. and L.A.G. wrote the manuscript. All authors discussed results and edited the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details accompanying the full-text HTML version of the paper at [www.nature.com/nature](http://www.nature.com/nature). Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to L.A.G. (Levi\_Garraway@dfci.harvard.edu).

## METHODS

**Center for Cancer Systems Biology (CCSB)/Broad Institute Kinase Open Reading Frame Collection.** We assembled a library of 597 kinase ORFs in pDONR-223 Entry vectors (Invitrogen). Individual clones were end-sequenced using vector-specific primers in both directions. Clones with substantial deviations from reported sequences were discarded. Entry clones and sequences are available via Addgene ([http://www.addgene.org/human\\_kinases](http://www.addgene.org/human_kinases)). Kinase ORFs were assembled from multiple sources; 337 kinases were isolated as single clones from the ORFeome 5.1 collection (<http://horfdb.dfci.harvard.edu>), 183 kinases were cloned from normal human tissue RNA (Ambion) by reverse transcription and subsequent PCR amplification to add Gateway sequences (Invitrogen), 64 kinases were cloned from templates provided by the Harvard Institute of Proteomics (HIP), and 13 kinases were cloned into the Gateway system from templates obtained from collaborating laboratories. The Gateway-compatible lentiviral vector pLX-Blast-V5 was created from the pLKO.1 backbone. LR Clonase enzymatic recombination reactions were performed to introduce the 597 kinases into pLX-Blast-V5 according to the manufacturer's protocol (Invitrogen).

**High throughput ORF screening.** A375 melanoma cells were plated in 384-well microtitre plates (500 cells per well). The following day, cells were spin-infected with the lentivirally-packaged kinase ORF library in the presence of  $8 \mu\text{g ml}^{-1}$  polybrene. At 48 h post-infection, media were replaced with standard growth media (two replicates), media containing  $1 \mu\text{M}$  PLX4720 (two replicates, two time points) or media containing  $10 \mu\text{g ml}^{-1}$  blasticidin (two replicates). After 4 and 6 days, cell growth was assayed using Cell Titer-Glo (Promega) according to manufacturer instructions. The entire experiment was performed twice.

**Identification of candidate resistance ORFs.** Raw luminescence values were imported into Microsoft Excel. Infection efficiency was determined by the percentage of duplicate-averaged raw luminescence in blasticidin-selected cells relative to non-selected cells. ORFs with an infection efficiency of less than 0.70 were excluded from further analysis along with any ORF having a standard deviation of  $>15,000$  raw luminescence units between duplicates. To identify ORFs whose expression affects proliferation, we compared the duplicate-averaged raw luminescence of individual ORFs against the average and standard deviation of all control-treated cells via the  $z$ -score, or standard score, below,

$$z = \frac{x - \mu}{\sigma}$$

where  $x$  is the average raw luminescence of a given ORF,  $\mu$  is the mean raw luminescence of all ORFs and  $\sigma$  is the standard deviation of the raw luminescence of all wells. Any individual ORF with a  $z$ -score  $> +2$  or  $< -2$  was annotated as affecting proliferation and removed from final analysis. Differential proliferation was determined by the percentage of duplicate-averaged raw luminescence values in PLX4720 ( $1 \mu\text{M}$ )-treated cells relative to untreated cells. Subsequently, differential proliferation was normalized to the positive control for PLX4720 resistance, MEK1(S218/222D) (MEK1<sup>DD</sup>), with MEK1<sup>DD</sup> differential proliferation = 1.0. MEK1<sup>DD</sup>-normalized differential proliferation for each individual ORF was averaged across two duplicate experiments, with two time points for each experiment (day 4 and day 6). A  $z$ -score was then generated, as described above, for average MEK1<sup>DD</sup>-normalized differential proliferation. ORFs with a  $z$ -score of  $>2$  were considered hits and were followed up in the secondary screen.

**ORF and shRNA expression.** ORFs were expressed from pLX-Blast-V5 (lentiviral) or pWZL-Blast, pBABE-Puro or pBABE-zeocin (retroviral) expression plasmids. For lentiviral transduction, 293T cells were transfected with  $1 \mu\text{g}$  of pLX-Blast-V5-ORF or pLKO.1-shRNA, 900 ng  $\Delta 8.9$  (*gag*, *pol*) and 100 ng VSV-G using 6  $\mu\text{l}$  Eugene6 transfection reagent (Roche). Viral supernatant was harvested 72 h post-transfection. Mammalian cells were infected at a 1:10–1:20 dilution of virus in 6-well plates in the presence of  $5 \mu\text{g ml}^{-1}$  polybrene and centrifuged at 2,250 r.p.m. (1,178g) for 1 h at  $37^\circ\text{C}$ . Twenty-four hours after infection blasticidin (pLX-Blast-V5,  $10 \mu\text{g ml}^{-1}$ ) or puromycin (pLKO.1,  $0.75 \mu\text{g ml}^{-1}$ ) was added and cells were selected for 48 h. For retrovirus production, 293T cells were transfected with  $1 \mu\text{g}$  of retroviral plasmid-ORF,  $1 \mu\text{g}$  pCL-AMPHO and 100 ng VSV-G, as described above. Cells were infected with retrovirus containing supernatant at a 1:2 dilution in  $5 \mu\text{g ml}^{-1}$  polybrene overnight, followed by media change to growth media. Infection was repeated once more (twice total), followed by selection, above.

**Secondary screen.** A375 ( $1.5 \times 10^3$ ) and SKMEL28 cells ( $3 \times 10^3$ ) were seeded in 96-well plates for 18 h. ORF-expressing lentivirus was added at a 1:10 dilution in the presence of  $8 \mu\text{g ml}^{-1}$  polybrene, and centrifuged at 2,250 r.p.m. (1,178g) and  $37^\circ\text{C}$  for 1 h. Following centrifugation, virus-containing media were changed to normal growth media and allowed to incubate for 18 h. Twenty-four hours after infection, DMSO (1:1,000) or  $10\times$  PLX4720 (in DMSO) was added to a final concentration of 100, 10, 1, 0.1, 0.01, 0.001, 0.0001 or 0.00001  $\mu\text{M}$ . Cell viability was assayed using WST-1 (Roche), per manufacturer recommendation, 4 days after the addition of PLX4720.

**Cell lines and reagents.** A375, SKMEL28, SKMEL30, COLO-679, WM451lu, SKMEL5, Malmé 3M, SKMEL30, WM3627, WM1976, WM3163, WM3130, WM3629, WM3453, WM3682 and WM3702 were all grown in RPMI (Cellgro), 10% FBS and 1% penicillin/streptomycin. M307 was grown in RPMI (Cellgro), 10% FBS and 1% penicillin/streptomycin supplemented with 1 mM sodium pyruvate. 293T, OUMS-23 and RPMI-7951 cells (ATCC) were grown in MEM (Cellgro), 10% FBS and 1% penicillin/streptomycin. Wild-type primary melanocytes were grown in HAM's F10 (Cellgro), 10% FBS and 1% penicillin/streptomycin. B-RAF(V600E)-expressing primary melanocytes were grown in TIVA media (Ham's F-10 (Cellgro), 7% FBS, 1% penicillin/streptomycin, 2 mM glutamine (Cellgro),  $100 \mu\text{M}$  IBMX,  $50 \text{ ng ml}^{-1}$  TPA (12-*O*-tetradecanoyl-phorbol-13-acetate), 1 mM 3',5'-cyclic AMP dibutyrate (dbcAMP; Sigma) and  $1 \mu\text{M}$  sodium vanadate). CI-1040 (PubChem ID: 6918454) was purchased from Shanghai Lechen International Trading Co., AZD6244 (PubChem ID: 10127622) from Selleck Chemicals, and PLX4720 (PubChem ID: 24180719) from Symansis. RAF265 (PubChem ID: 11656518) was a generous gift from Novartis Pharma AG. Unless otherwise indicated, all drug treatments were for 16 h. Activated alleles of NRAS and KRAS have been described previously<sup>28,29</sup>.

**Pharmacologic growth inhibition assays.** Cultured cells were seeded into 96-well plates (3,000 cells per well) for all melanoma cell lines; 1,500 cells were seeded for A375. Twenty-four hours after seeding, serial dilutions of the relevant compound were prepared in DMSO added to cells, yielding final drug concentrations ranging from  $100 \mu\text{M}$  to  $1 \times 10^{-5} \mu\text{M}$ , with the final volume of DMSO not exceeding 1%. Cells were incubated for 96 h following addition of drug. Cell viability was measured using the WST1 viability assay (Roche). Viability was calculated as a percentage of control (untreated cells) after background subtraction. A minimum of six replicates were performed for each cell line and drug combination. Data from growth-inhibition assays were modelled using a nonlinear regression curve fit with a sigmoid dose-response. These curves were displayed and  $\text{GI}_{50}$  generated using GraphPad Prism 5 for Windows (GraphPad). Sigmoid-response curves that crossed the 50% inhibition point at or above  $10 \mu\text{M}$  have  $\text{GI}_{50}$  values annotated as  $>10 \mu\text{M}$ . For single-dose studies, the identical protocol was followed, using a single dose of indicated drug ( $1 \mu\text{M}$  unless otherwise noted).

**Immunoblots and immunoprecipitations.** Cells were washed twice with ice-cold PBS and lysed with 1% NP-40 buffer (150 mM NaCl, 50 mM Tris pH 7.5, 2 mM EDTA pH 8, 25 mM NaF and 1% NP-40) containing  $2\times$  protease inhibitors (Roche) and  $1\times$  Phosphatase Inhibitor Cocktails I and II (CalBioChem). Lysates were quantified (Bradford assay), normalized, reduced, denatured ( $95^\circ\text{C}$ ) and resolved by SDS gel electrophoresis on 10% Tris/Glycine gels (Invitrogen). Protein was transferred to PVDF membranes and probed with primary antibodies recognizing pERK1/2 (T202/Y204), pMEK1/2 (S217/221), MEK1/2, MEK1, MEK2, C-RAF (rabbit host), pC-RAF (pS338) (Cell Signaling Technology; 1:1,000), V5-HRP (HRP, horseradish peroxidase; Invitrogen; 1:5,000), COT (1:500), B-RAF (1:2,000), Actin (1:1,000), Actin-HRP (1:1,000; Santa Cruz), C-RAF (mouse host; 1:1,000; BD Transduction Labs), Vinculin (Sigma; 1:20,000), AXL (1:500; R&D Systems). After incubation with the appropriate secondary antibody (anti-rabbit, anti-mouse IgG, HRP-linked; 1:1,000 dilution, Cell Signaling Technology or anti-goat IgG, HRP-linked; 1:1,000 dilution; Santa Cruz), proteins were detected using chemiluminescence (Pierce). Immunoprecipitations were performed overnight at  $4^\circ\text{C}$  in 1% NP-40 lysis buffer, as described above, at a concentration of  $1 \mu\text{g ml}^{-1}$  total protein using an antibody recognizing C-RAF (1:50; Cell Signaling Technology). Antibody: antigen complexes were bound to Protein A agarose (25  $\mu\text{l}$ , 50% slurry; Pierce) for 2 h at  $4^\circ\text{C}$ . Beads were centrifuged and washed three times in lysis buffer and eluted and denatured ( $95^\circ\text{C}$ ) in  $2\times$  reduced sample buffer (Invitrogen). Immunoblots were performed as above. Phospho-protein quantification was performed using NIH ImageJ.

Lysates from tumour and matched normal skin were generated by mechanical homogenization of tissue in RIPA (50 mM Tris (pH 7.4), 150 mM NaCl, 1 mM EDTA, 0.1% SDS, 1.0% NaDOC (sodium deoxycholate), 1.0% Triton X-100, 25 mM NaF, 1 mM  $\text{Na}_3\text{VO}_4$ ) containing protease and phosphatase inhibitors, as above. Subsequent normalization and immunoblots were performed as above.

**Biopsied melanoma tumour material.** Biopsied tumour material consisted of discarded and de-identified tissue that was obtained with informed consent and characterized under protocol 02-017 (paired samples, Massachusetts General Hospital) and 07-087 (unpaired sample, Dana-Farber Cancer Institute). For paired specimens, 'on-treatment' samples were collected 10–14 days after initiation of PLX4032 treatment (Supplementary Table 4).

**Inhibition of COT kinase activity.** Adherent RPMI-7951 cells were washed twice with  $1\times$  PBS and incubated overnight in serum-free growth media. Subsequently, 4-(3-chloro-4-fluorophenylamino)-6-(pyridin-3-yl-methylamino)-3-cyano-[1,7]-naphthyridine (EMD; TPL2 inhibitor I; catalogue number 616373, PubChem ID:

9549300), suspended in DMSO at the indicated concentration, was added to cells for 1 h, after which protein extracts were made as described above.

**Quantitative RT-PCR.** mRNA was extracted from cell lines and fresh-frozen tumours using the RNeasy kit (Qiagen). Total mRNA was used for subsequent reverse transcription using the SuperScript III First-Strand Synthesis SuperMix (Invitrogen) for cell lines and unpaired tumour samples, and the SuperScript VILO cDNA synthesis kit (Invitrogen) for paired frozen tumour samples. The reverse transcription reaction (5 µl) was used for quantitative PCR using SYBR Green PCR Master Mix and gene-specific primers, in triplicate, using an ABI 7300 Real Time PCR System. Primers used for detection are as follows; *COT* forward: 5'-CAAGTGAAGAGCCAGCAGTTT-3'; *COT* reverse: 5'-GCAAGCAAATCC TCCACAGTTC-3'; *TBP* forward: 5'-CCCGAAACGCCGAATATAATCC-3'; *TBP* reverse: 5'-GACTGTTCTTCACTCTTGGCTC-3'; *GAPDH* forward: 5'-CATCATCTCTGCCCCCTCT-3'; *GAPDH* reverse: 5'-GGTGCTAAGCAG TTGGTGGT-3'.

**In vitro kinase assay.** *In vitro* kinase assays were performed as described previously<sup>15</sup> using 1 µg each of COT (amino acids 30–397, R&D Systems) and inactive ERK1 (Millipore).

**Cellular viability assays.** Adherent RPMI-7951 cells were infected with virus expressing shRNAs against *COT* or *Luciferase* as described above. Following selection, cells were plated ( $1.5 \times 10^5$  cells per well) onto a 24-well plate in quadruplicate. Viable cells were counted via trypan blue exclusion using a VI-CELL Cell Viability Analyser, following manufacturer's specifications. Quadruplicate cell counts were averaged and normalized relative to that of the control shRNA.

**The Cancer Cell Line Encyclopedia (CCLE).** The Cancer Cell Line Encyclopedia (CCLE) project is a collaboration between the Broad Institute, the Novartis Institutes for Biomedical Research (NIBR) and the Genomics Institute of the Novartis Research Foundation (GNF) to conduct a detailed genetic and pharmacologic characterization of a large panel of human cancer models, to develop

integrated computational analyses that link distinct pharmacologic vulnerabilities to genomic patterns and to translate cell line integrative genomics into cancer patient stratification. Chromosomal copy number and gene expression data used for this study are available online at <http://www.broadinstitute.org/cgi-bin/cancer/datasets.cgi>.

**Expression profiling of cancer cell lines.** We carried out oligonucleotide microarray analysis using the GeneChip Human Genome U133 Plus 2.0 Affymetrix expression array (Affymetrix). Samples were converted to labelled, fragmented, cRNA following the Affymetrix protocol for use on the expression microarray.

**shRNA constructs used (pLKO.1).** The shRNA constructs used were shLuc (TRCN0000072243, 5'-CTTCGAAATGTCCGTTCCGTT-3'), shBRAF(1) (TRCN0000006289, NM\_004333.2-1106s1c1, 5'-CTTCGAAATGTCCGTTCCGTT-3'), shBRAF(2) (TRCN0000006291, NM\_004333.2-2267s1c1, 5'-GCTGG TTTCCAAACAGAGGAT-3'), shCRAF(1) (TRCN0000001066, NM\_002880.x-1236s1c1, 5'-CGGAGATGTTGCAGTAAAGAT-3'), shCRAF(2) (TRCN0000001068, NM\_002880.x-1529s1c1, 5'-GAGACATGAAATCCAACAATA-3'), shMEK1(1) (TRCN0000002332, NM\_002755.x-1015s1c1, 5'-GATTACATAGTCAACGAG CCT-3'), shMEK1(2) (TRCN0000002329, NM\_002755.x-455s1c1, 5'-GCTTC TATGGTGCGTTCTACA-3'), shMEK2(1) (TRCN0000007007, NM\_030662.2-1219s1c1, 5'-TGGACTATATTGTGAACGAGC-3'), shMEK2(2) (TRCN0000007005, NM\_030662.2-847s1c1, 5'-CCAACATCCTCGTGAACCTA-3'), shCOT(1) (TRCN0000010013, NM\_005204.x-1826s1c1, 5'-CAAGAGCCGCAGACCTAC TAA-3') and shCOT(2) (TRCN0000196518, NM\_005204.2-2809s1c1, 5'-GAT GAGAATGTGACCTTTAAG-3').

28. Boehm, J. S. *et al.* Integrative genomic approaches identify *IKBKE* as a breast cancer oncogene. *Cell* **129**, 1065–1079 (2007).
29. Lundberg, A. S. *et al.* Immortalization and transformation of primary human airway epithelial cells by gene transfer. *Oncogene* **21**, 4577–4586 (2002).



# Melanomas acquire resistance to B-RAF(V600E) inhibition by RTK or N-RAS upregulation

Ramin Nazarian<sup>1,2\*</sup>, Hubing Shi<sup>1,2\*</sup>, Qi Wang<sup>1,2</sup>, Xiangju Kong<sup>1,2</sup>, Richard C. Koya<sup>2,3</sup>, Hane Lee<sup>2,4</sup>, Zugen Chen<sup>2,4</sup>, Mi-Kyung Lee<sup>1,2</sup>, Narsis Attar<sup>2,5</sup>, Hooman Sazegar<sup>2,5</sup>, Thinnie Chodon<sup>2,5</sup>, Stanley F. Nelson<sup>2,4,6</sup>, Grant McArthur<sup>7</sup>, Jeffrey A. Sosman<sup>8</sup>, Antoni Ribas<sup>2,3,5</sup> & Roger S. Lo<sup>1,2</sup>

Activating B-RAF(V600E) (also known as BRAF) kinase mutations occur in ~7% of human malignancies and ~60% of melanomas<sup>1</sup>. Early clinical experience with a novel class I RAF-selective inhibitor, PLX4032, demonstrated an unprecedented 80% anti-tumour response rate among patients with B-RAF(V600E)-positive melanomas, but acquired drug resistance frequently develops after initial responses<sup>2</sup>. Hypotheses for mechanisms of acquired resistance to B-RAF inhibition include secondary mutations in *B-RAF(V600E)*, MAPK reactivation, and activation of alternative survival pathways<sup>3–5</sup>. Here we show that acquired resistance to PLX4032 develops by mutually exclusive PDGFR $\beta$  (also known as PDGFRB) upregulation or *N-RAS* (also known as *NRAS*) mutations but not through secondary mutations in *B-RAF(V600E)*. We used PLX4032-resistant sub-lines artificially derived from *B-RAF(V600E)*-positive melanoma cell lines and validated key findings in PLX4032-resistant tumours and tumour-matched, short-term cultures from clinical trial patients. Induction of PDGFR $\beta$  RNA, protein and tyrosine phosphorylation emerged as a dominant feature of acquired PLX4032 resistance in a subset of melanoma sub-lines, patient-derived biopsies and short-term cultures. PDGFR $\beta$ -upregulated tumour cells have low activated RAS levels and, when treated with PLX4032, do not reactivate the MAPK pathway significantly. In another subset, high levels of activated N-RAS resulting from mutations lead to significant MAPK pathway reactivation upon PLX4032 treatment. Knockdown of *PDGFR $\beta$*  or *N-RAS* reduced growth of the respective PLX4032-resistant subsets. Overexpression of PDGFR $\beta$  or N-RAS(Q61K) conferred PLX4032 resistance to PLX4032-sensitive parental cell lines. Importantly, MAPK reactivation predicts MEK inhibitor sensitivity. Thus, melanomas escape B-RAF(V600E) targeting not through secondary B-RAF(V600E) mutations but via receptor tyrosine kinase (RTK)-mediated activation of alternative survival pathway(s) or activated RAS-mediated reactivation of the MAPK pathway, suggesting additional therapeutic strategies.

We selected three *B-RAF(V600E)*-positive parental (P) cell lines, M229, M238 and M249, exquisitely sensitive to PLX4032-mediated growth inhibition *in vitro* and *in vivo*<sup>6</sup>, and derived PLX4032-resistant (R) sub-lines by chronic PLX4032 exposure. In cell survival assays, M229 R, M238 R and M249 R sub-lines displayed strong resistance to PLX4032 (GI<sub>50</sub>, the concentration of drug that inhibits growth of cells by 50%, not reached up to 10  $\mu$ M) and paradoxically enhanced growth at low PLX4032 concentrations, in contrast to parental cells (Supplementary Fig. 1a). Morphologically, both M229 R and M238 R sub-lines appear flatter and more fibroblast-like compared to their parental counterparts, but this morphologic switch was not seen in the M249 P versus M249 R4 pair (Supplementary Fig. 2a).

There were no secondary mutations in the drug target *B-RAF(V600E)* observed on bi-directional Sanger sequencing of all 18 *B-RAF* exons in 15 M229 R (R1–R15), two M238 R (R1 and R2), and one M249 R (R4) acquired resistant sub-lines (Supplementary Table 1 and Supplementary Fig. 3a, left column). Based on Sanger sequencing, this lack of secondary *B-RAF(V600E)* mutation along with retention of the original *B-RAF(V600E)* mutation was confirmed in 16/16 melanoma tumour biopsies (from 12 patients) with clinically acquired resistance to PLX4032 (that is, initial >30% tumour size decrease or partial response, as defined by RECIST (response evaluation criteria in solid tumours) and subsequent progression on PLX4032 dosing; see examples in Supplementary Fig. 4) and 5/5 short-term melanoma cultures established from 5 resistant tumours obtained from 4 patients (Supplementary Table 2). Given recent reports of B-RAF-selective inhibitors having a growth-promoting effect on *B-RAF* wild-type tumour cells<sup>7–9</sup>, retention of the original *B-RAF* alleles in PLX4032-resistant sub-lines, tissues and cultures indicates that PLX4032 chronic treatment did not select for the outgrowth of a pre-existing, minor *B-RAF* wild-type sub-population. Furthermore, immunoprecipitated B-RAF kinase activities from resistant sub-lines and short-term cultures were similarly sensitive to PLX4032 as B-RAF kinase activities immunoprecipitated from parental cell lines (Supplementary Fig. 3b; Pt48 R and Pt55 R resistance to PLX4032 (ref. 10) and the pre-clinical analogue PLX4720 (ref. 11) shown in Supplementary Fig. 5a and b, respectively; Pt, patient). These results demonstrate that, in all tested acquired resistant cell lines and cultures, the mutated B-RAF(V600E) kinase lacks secondary mutations and hence retains its ability to respond to PLX4032.

Given that minority PLX4032-resistant sub-populations in tissues may acquire B-RAF(V600E) secondary mutations not detectable by Sanger sequencing, we analysed “ultra-deep” (Supplementary Fig. 6) and deep (Supplementary Fig. 7) sequences of *B-RAF* (exons 2–18) using the Illumina platform for 9/11 acquired resistant tumour samples without tumour-matched short-term cultures (one sample, Pt111-010 DP2, intentionally analysed by both methods; DP, disease progression). Ultradeep *B-RAF* sequencing of five PLX4032-resistant melanoma tissues resulted in every base of exons 2–18 being sequenced at a median coverage of 127 $\times$  (27 $\times$ –128 $\times$ ) (Supplementary Fig. 6a and b). The known variant, V600E, was detected in all five samples with significantly high non-reference allele frequencies (NAF) (Supplementary Fig. 6c). In all five tissues, exon 13, where the T529 gatekeeper residue<sup>12</sup> is located, was independently amplified and uniquely bar-coded twice. Rare variants (none at the T529 codon; Supplementary Fig. 6d) detected in these independent exon 13 analyses do not overlap and helped define the true, signal NAF at >4.81% (Supplementary

<sup>1</sup>Division of Dermatology/Department of Medicine, UCLA's Jonsson Comprehensive Cancer Center, 52-121 CHS, Los Angeles, California 90095-1750, USA. <sup>2</sup>David Geffen School of Medicine, University of California, Los Angeles, California 90095-1750, USA. <sup>3</sup>Division of Surgical Oncology/Department of Surgery, UCLA's Jonsson Comprehensive Cancer Center, 54-140 CHS, Los Angeles, California 90095-1782, USA. <sup>4</sup>Department of Human Genetics, UCLA's Jonsson Comprehensive Cancer Center, 5506A Gonda Center, Los Angeles, California 90095-7088, USA. <sup>5</sup>Division of Hematology & Oncology/Department of Medicine, UCLA's Jonsson Comprehensive Cancer Center, 9-954 Factor Building, Los Angeles, California 90095-1678, USA. <sup>6</sup>Department of Pediatrics, UCLA's Jonsson Comprehensive Cancer Center, 5506A Gonda Center, Los Angeles, California 90095-7088, USA. <sup>7</sup>Peter MacCallum Cancer Center, St Andrews Place, East Melbourne 3002, Australia. <sup>8</sup>Department of Medicine, Vanderbilt-Ingram Cancer Center, 777 Preston Research Building, Nashville, Tennessee 37232-6838, USA.

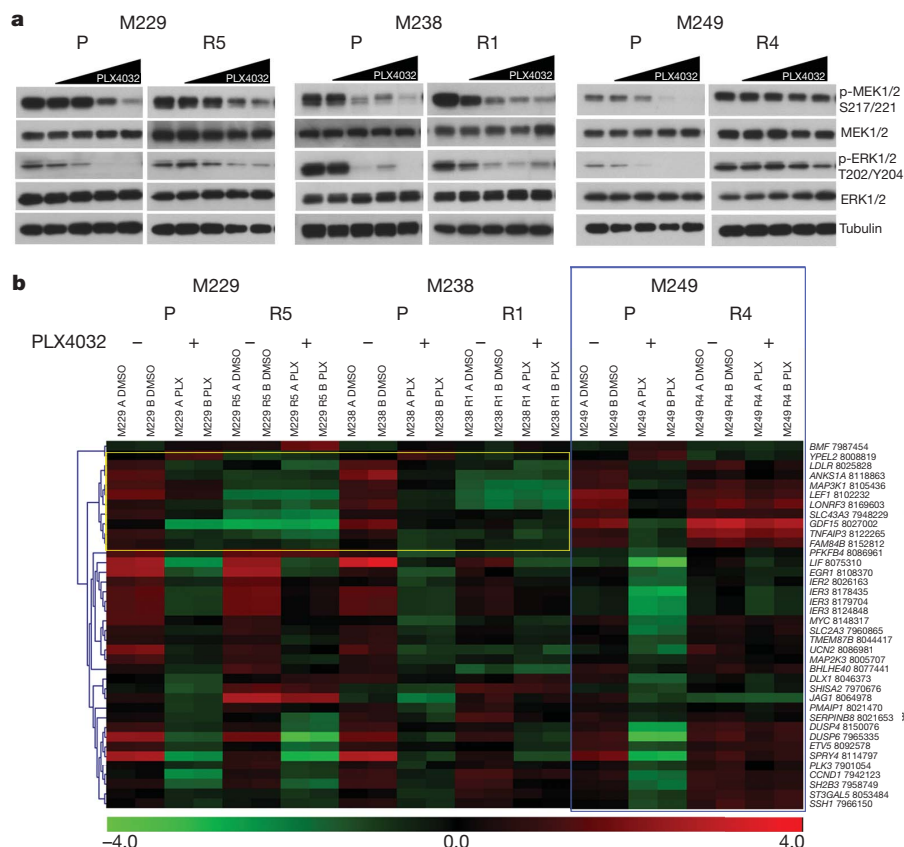
\*These authors contributed equally to this work.

Methods). Furthermore, deep *B-RAF* (exons 2–18) sequence analysis of PLX4032-resistant melanoma tissues from a whole exome sequencing project resulted in 2,396 base pairs of *B-RAF* coding regions having coverage  $\geq 10\times$  (average coverage per exon in each tissue shown in Supplementary Fig. 7a). After filtering, no position harboured a variant with a NAF  $>4.81\%$ , except for the known V600E mutation in all five resistant samples. Together, these data strongly corroborate the lack of *B-RAF*(V600E) secondary mutations during the evolution of PLX4032 acquired resistance in the majority of patients and their tumours.

To begin to understand PLX4032-resistance *in vitro*, we used phospho-specific antibodies to probe the activation status of the RAF downstream effectors, MEK1/2 and ERK1/2 (also known as MAP2K1/2 and MAPK3/1, respectively), in parental versus resistant sub-lines, with and without PLX4032 (Fig. 1a). As expected, PLX4032 induced dose-dependent decreases in p-MEK1/2 and p-ERK1/2 in all parental cells. However, the pattern of MEK-ERK sensitivity to PLX4032 varied among resistant sub-lines, suggesting distinct mechanisms. In contrast to M249 R4, which showed strong resistance to PLX4032-induced MEK/ERK inhibition (suggesting MAPK reactivation), M229 R5 and M238 R1 were both similarly sensitive to PLX4032-induced decreases in the levels of p-MEK1/2 and p-ERK1/2. Gene expression profiling (Fig. 1b) further supported distinct PLX4032 acquired resistant mechanisms represented by M229 R5/M238 R1 versus M249 R4. We first used the gene expression alterations responsive to PLX4032 in parental cells to define a *B-RAF*(V600E)-responsive gene signature, which is similar to gene sets defined by a MEK1 inhibitor (PD325901)<sup>13</sup> and by PLX4720 (ref. 14; Supplementary Fig. 8). Concordant with the western blot results (Fig. 1a), M249 R4 demonstrated striking resistance to PLX4032 treatment with a gene signature of persistent MEK-ERK activation, whereas both M229 R5 and M238 R1 retained a PLX4032-sensitive gene signature (Fig. 1b). These data confirm that M229 R5 and M238 R1 share key characteristics of resistance, which are in line with unsupervised clustering of these two resistant sub-lines in genome-wide, differential expression patterns (Supplementary Fig. 9).

Gene set enrichment analysis demonstrated an enrichment of RTK-controlled signalling in M229 R5 and M238 R1 but exclusive of M249 R4 (Supplementary Table 3). Unsupervised clustering of the receptor tyrosine kinase gene expression profiles showed that M229 R5 and M238 R1 clustered away from M229 and M238 parental cell lines largely based on higher expression levels of *KIT*, *MET*, *EGFR* and *PDGFR $\beta$*  (Supplementary Fig. 10a, yellow highlight). RNA upregulation of these four RTKs was consistently not associated with genomic DNA (gDNA) copy number gain (Supplementary Fig. 10b). Of these four candidate RTKs, EGFR and PDGFR $\beta$  protein levels were overexpressed (Fig. 2a, left; Fig. 3b; Supplementary Fig. 10c), but only PDGFR $\beta$  displayed elevated activation-associated tyrosine phosphorylation in a phospho-RTK array (Fig. 2a, right). PDGFR $\beta$  RNA upregulation was a common feature among additional M229 R and M238 R sub-lines (Supplementary Fig. 11a) but could not be observed in any of ten randomly selected parental melanoma cell lines (Supplementary Fig. 11b). Interestingly, tyrosine phosphorylation of PDGFR $\beta$  correlated with an upregulation of a gene signature unique to PDGFR $\beta$  (ref. 15; Supplementary Fig. 12) but is not due to mutational activation, as PDGFR $\beta$  cDNAs derived from M229 R5, M238 R1 and Pt48 R are wild type (Supplementary Table 1).

We then validated our *in vitro* finding *in vivo* by studying clinical trial patient-derived samples (Supplementary Table 2; Fig. 2b) and tumour-matched short-term cultures (Fig. 2c and d). In 4/11 available, paired biopsy specimens, the resistant tumours showed a tumour-associated overexpression of PDGFR $\beta$  compared to the baseline tumour in the same patients (Fig. 2b; Supplementary Table 2 and Supplementary Fig. 13). PDGFR $\beta$ -positive areas of tissue sections were consistently strongly positive for S100 or MART1 (melanoma markers; MART1 is also known as MLANA) but lacked CD31 (an endothelial, platelet, macrophage marker, also known as PECAM1) staining (data not shown). We were able to validate this finding further in an available short-term culture (Pt48 R) derived from a PLX4032-resistant, PDGFR $\beta$ -positive tumour. Pt48 R was established from an intracardiac mass progressing 6 months after initiating treatment with PLX4032.

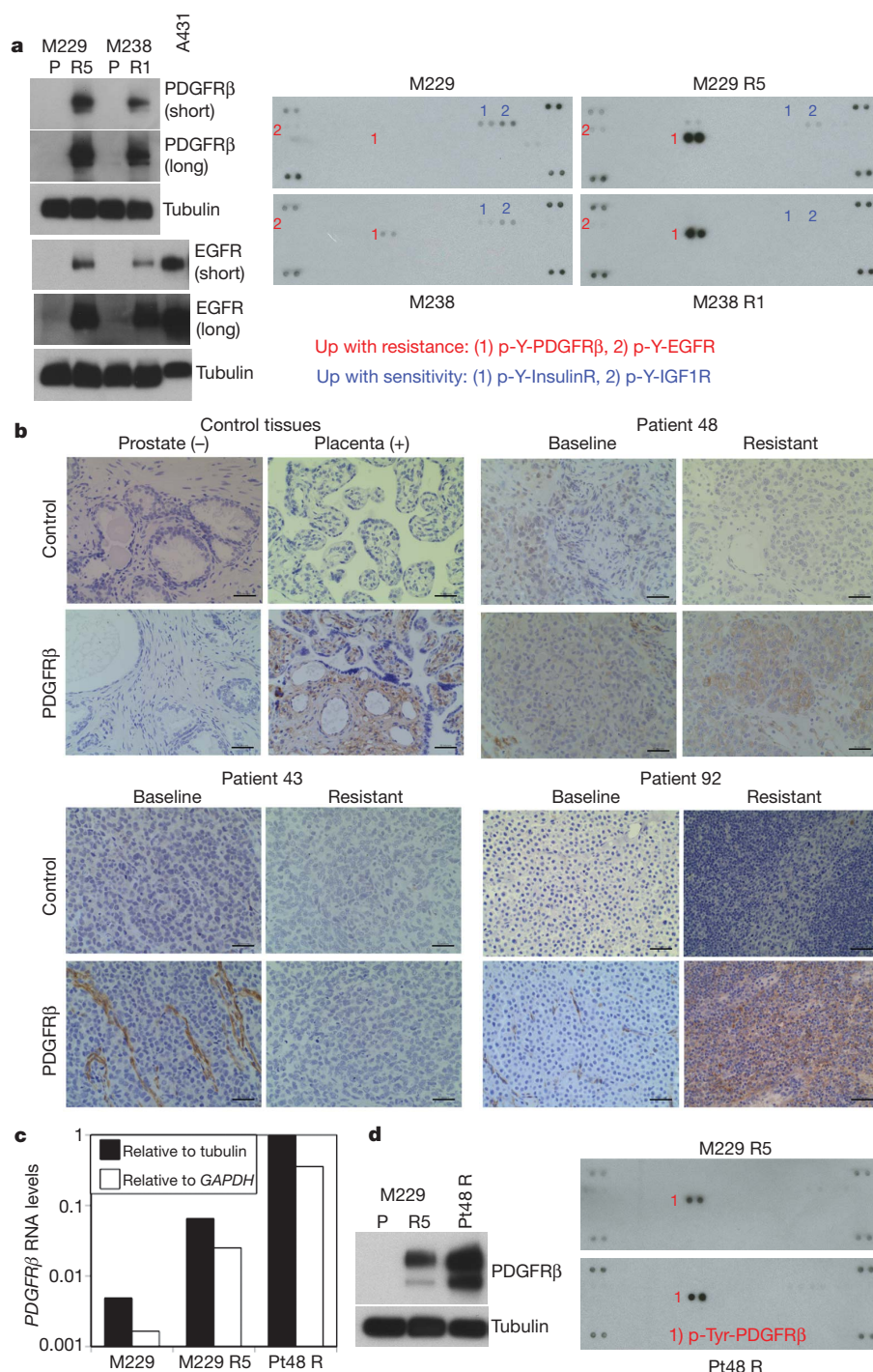


**Figure 1 | *In vitro* models of PLX4032 acquired resistance display differential MAPK reactivation.** **a**, Parental and PLX4032-resistant sub-lines were treated with increasing PLX4032 concentration (0, 0.01, 0.1, 1 and 10  $\mu$ M), and the effects on MAPK signalling were determined by immunoblotting for p-MEK1/2 and p-ERK1/2 levels. Total MEK1/2, ERK1/2 and tubulin levels, loading controls. **b**, Heat map for B-RAF(V600E) signature genes in each of the cell lines treated with DMSO or PLX4032. Colour scale, log<sub>2</sub>-transformed expression (red, high; green, low) for each gene (row) normalized by the mean of all samples. Blue box showing M249 R4 MAPK reactivation. Yellow box showing diminished, baseline expression of B-RAF(V600E) signature genes in M229 and M238 resistant sub-lines (FDR < 0.05). The probeset number is shown after each gene.



**Figure 2 | PDGFR $\beta$  upregulation is strongly correlated with PLX4032 acquired resistance.**

**a**, Left, total levels of PDGFR $\beta$  and EGFR. A431, an EGFR-amplified cell line. Tubulin levels, loading control. Right, whole-cell extracts were incubated on the RTK antibody arrays, and phosphorylation status was determined by subsequent incubation with anti-phosphotyrosine horseradish peroxidase (each RTK spotted in duplicate, positive controls in corners, gene identity below). **b**, Anti-PDGFR $\beta$  immunohistochemistry of formalin-fixed, paraffin-embedded tissues. Prostate, negative control; placenta, positive control. Black bar, 50  $\mu$ m. **c**, Relative RNA levels of PDGFR $\beta$  in M229 P/R5 and Pt48 R as determined by real-time, quantitative PCR (average of duplicates). **d**, Total PDGFR $\beta$  (left) and p-RTK (right) levels in Pt48 R versus M229 R5.

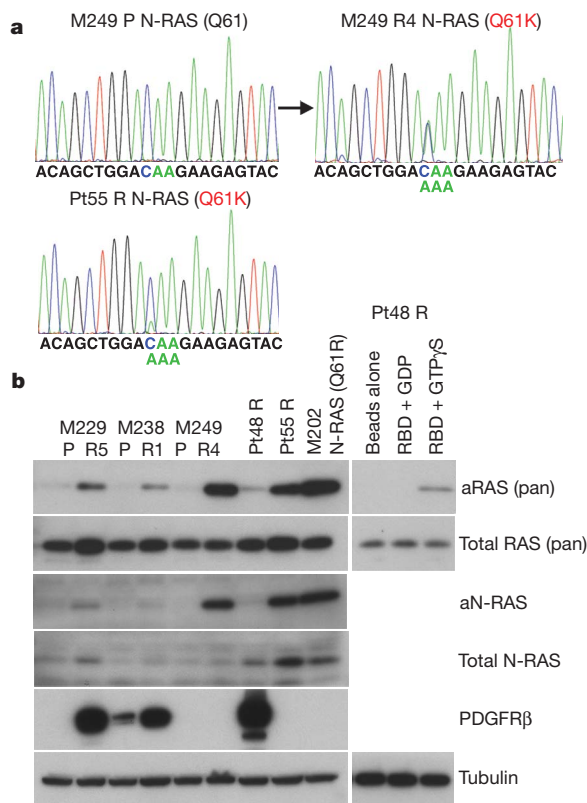


The Pt48 R short-term culture demonstrated clear overexpression of PDGFR $\beta$  RNA (Fig. 2c), protein and p-Tyr levels (Fig. 2d).

In M249 R4, we sequenced all exons of *N-RAS*, *K-RAS* (also known as *KRAS*) or *H-RAS* (also known as *HRAS*) (to include codons 12, 13, and 61 as well as mutational hotspots of emerging significance<sup>16</sup>) and *MEK1* (ref. 17; Supplementary Table 1 and data not shown) because we proposed a resistance mechanism reactivating MAPK despite not having a secondary *B-RAF* mutation. Interestingly, M249 R4 harbours a *N-RAS*(Q61K) activating mutation not present in the parental M249 cell line (Fig. 3a). We found *N-RAS* mutations in 2/16 acquired resistant biopsy samples (note that both came from Pt55; Supplementary Table 2 and Supplementary Fig. 14). A *N-RAS*(Q61K) mutated sample, Pt55 DP1 (for disease progression 1) was obtained from a biopsy taken from an isolated, nodal metastasis that partially regressed on PLX4032 but increased in size 10 months after starting on therapy with PLX4032

(Supplementary Fig. 4a). This patient continued on therapy with PLX4032 until 6 months later, when several other nodal metastases developed (Supplementary Fig. 14a, b). Analysis of a biopsy taken at a second progression site (Pt55 DP2) demonstrated a different mutation in *N-RAS*, *N-RAS*(Q61R) (Supplementary Fig. 14b). Both Pt55 DP1 and DP2 tissue *N-RAS* mutations were confirmed in their respective short-term cultures, Pt55 R and Pt55 R2 (Fig. 3a and Supplementary Fig. 14b). Also, both DP1 and DP2 (and their respective cultures) harboured increased *N-RAS* gDNA copy numbers (Supplementary Fig. 14c and d). Both Pt55 R and Pt55 R2 also showed increased *N-RAS* RNA (Supplementary Fig. 14e) and protein levels (Fig. 3b). In addition, *N-RAS*(Q61K) mutation in M249 R4 and Pt55 R correlated with a marked increase in activated *N-RAS* levels (Fig. 3b). Of note, the *N-RAS* mutations were mutually exclusive with PDGFR $\beta$  overexpression in all samples (Supplementary Table 2).



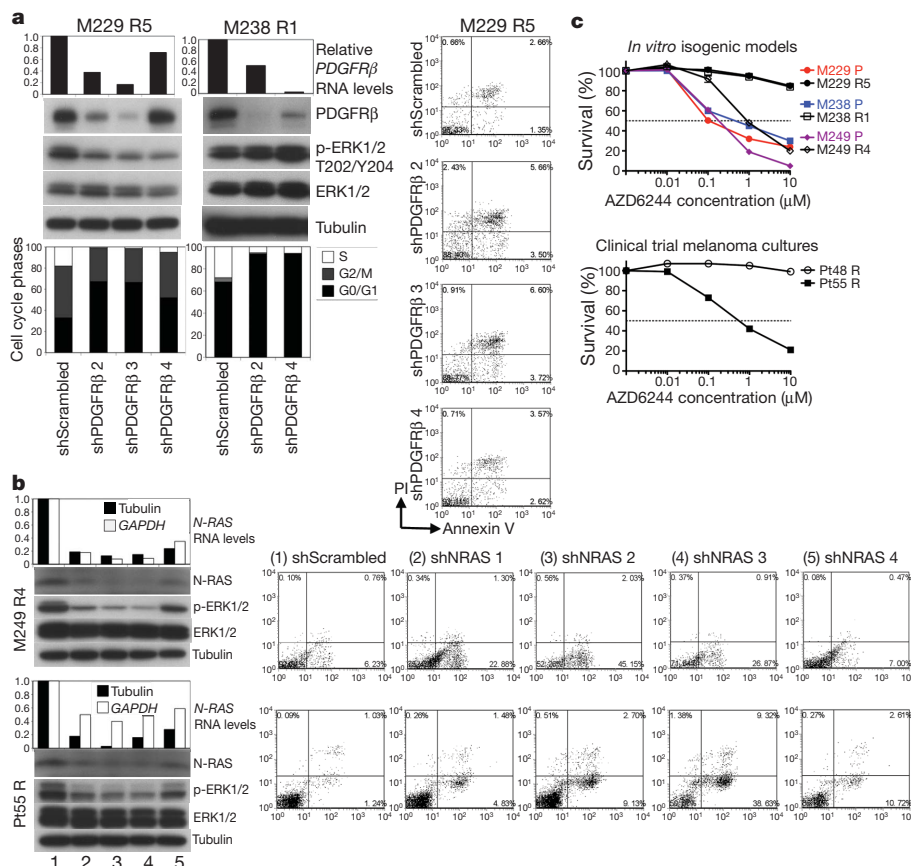


Knockdown of *PDGFRβ* or *N-RAS* using small interfering RNA (siRNA) pools preferentially growth-inhibited melanoma cells with upregulated *PDGFRβ* or *N-RAS*, respectively (Supplementary Fig. 15a, b and Supplementary Table 4). We then selected two resistant sub-lines

**Figure 3 | N-RAS upregulation correlates with a distinct subset of PLX4032 acquired resistance.** **a**, Detection of a *N-RAS*(Q61K) allele in M249 R4 and Pt55 R. **b**, The levels of activated RAS (aRAS) and N-RAS (aN-RAS) eluted after pull-down using the RAS-binding domain (RBD) of RAF-1. The total levels of RAS, N-RAS, *PDGFRβ* and tubulin (loading control) from the whole-cell lysates are shown by immunoblotting. Effects of GDP and GTPγS pre-incubation on RBD pull-down and beads without RBD pull-down from Pt48 R lysates are shown as controls.

or cultures to test the effects of individual *PDGFRβ* and *N-RAS* short hairpin RNAs (shRNAs; Fig. 4a and b, respectively). Stable knockdown of *PDGFRβ* caused an admixture of G0/G1 cell cycle arrest (in a MEK inhibitor-dependent manner due to compensatory signalling; Supplementary Fig. 16a and data not shown) and apoptosis in M229 R5 and a G0/G1 cell cycle arrest in M238 R1. This effect was specific, as stable *PDGFRβ* knockdown in M249 R4 and Pt55 R did not result in G0/G1 cell cycle arrest (Supplementary Fig. 17a). In contrast, stable *N-RAS* knockdown resulted in a predominantly apoptotic response in M249 R4 and Pt55 R (Fig. 4b) but not in M229 R5, M238 R1 or Pt48 R (Supplementary Fig. 17b). Moreover, stable *N-RAS* knockdown markedly conferred PLX4032 sensitivity to M249 R4 and Pt55 R but had no effect on M229 R5 PLX4032 resistance (Supplementary Fig. 18a). Flag-*N-RAS*(Q61K) stable overexpression conferred PLX4032 resistance in the M249 parental cell line (Supplementary Fig. 18b), whereas stable *PDGFRβ*-MYC overexpression conferred reduced PLX4032 sensitivity in both M229 and M238 parental cell lines (Supplementary Fig. 19).

We then asked whether *N-RAS*-dependent growth and reactivation of the MAPK pathway (Fig. 1a and Supplementary Fig. 20) would selectively sensitize M249 R4 and Pt55 R to MEK inhibition. Indeed, whereas the growth of M229 R5, M238 R1 and Pt48 R was uniformly highly resistant to the MEK inhibitor AZD6244 (and U0126, Supplementary Fig. 21), the growth of M249 R4 and Pt55 R was sensitive to MEK inhibition in the presence of PLX4032 (Fig. 4c) or absence of PLX4032 (Supplementary Fig. 22). It is known that activated *N-RAS* in melanoma cells uses C-RAF (also known as RAF1) over B-RAF to



**Figure 4 | *PDGFRβ*- and *N-RAS*-mediated growth and survival pathways differentially predict MEK inhibitor sensitivity.**

**a**, Transduction of *PDGFRβ* shRNAs in M229 R5 and M238 R1 (1 μM PLX4032), RNA (relative to *GAPDH*) and protein knockdown, effects on p-ERK levels, cell cycle distribution, and apoptosis (when applicable). M229 R5 was also treated with 0.5 μM AZD6244. PI, propidium iodide. **b**, Transduction of *N-RAS* shRNAs in M249 R4 and Pt55 R (1 μM PLX4032), RNA and protein knockdown, effects on p-ERK levels and apoptosis. **c**, Survival curves for isogenic cell line pairs and melanoma cultures treated with the indicated AZD6244 concentration for 72 h (relative to DMSO-treated controls; mean ± s.e.m.,  $n = 5$ ). PLX4032-resistant cells were grown with PLX4032. Dashed line, 50% cell killing.

signal to MEK-ERK<sup>18</sup>. Thus, N-RAS activation would be capable of bypassing PLX4032-inhibited B-RAF, reactivating the MAPK pathway. It is worth noting that PDGFR $\beta$ -upregulated, PLX4032-resistant melanoma sub-lines (M229 R5 and M238 R1) and culture (Pt48 R) are resistant not only to AZD6244 but also to imatinib, which is at least partially due to rebound, compensatory survival signalling (Supplementary Fig. 23 and unpublished observations, H.S. and R.S.L.).

We propose (Supplementary Fig. 1b) that *B-RAF(V600E)*-positive melanomas, instead of accumulating *B-RAF(V600E)* secondary mutations, can acquire PLX4032 resistance by (1) activating an RTK (PDGFR $\beta$ )-dependent survival pathway in addition to MAPK, or (2) reactivating the MAPK pathway via N-RAS upregulation. These two mechanisms account for acquired PLX4032 resistance in 5/12 patients in our study cohort, and additional mechanisms await future discovery. Some patients who relapse on PLX4032 are already being enrolled in a phase II MEK inhibitor trial (ClinicalTrials.gov identifier NCT01037127) based on the assumption of MAPK reactivation. Our findings directly imply a strategy to stratify patients who relapse on PLX4032 and should prompt a search for rational combinations of targeting agents most optimal for distinct mechanisms of acquired resistance to PLX4032 as well as other B-RAF inhibitors (for example, GSK2118436) in clinical development.

## METHODS SUMMARY

**Cell culture, infections and compounds.** Cells were maintained in Dulbecco's modified Eagle medium (DMEM) with 10 or 20% fetal bovine serum and glutamine. shRNAs were sub-cloned into the lentiviral vector pLL3.7 and infections carried out with protamine sulphate. Stocks of PLX4032 (Plexxikon) and AZD6244 (commercially available) were made in DMSO. Cells were quantified using CellTiter-GLO Luminescence (Promega).

**Protein detection.** Western blots were probed with antibodies against p-MEK1/2 (S217/221), MEK1/2, p-ERK1/2 (T202/Y204), ERK1/2, PDGFR $\beta$ , and EGFR (Cell Signaling Technologies), and N-RAS (Santa Cruz Biotechnology), pan-RAS (Thermo Scientific) and tubulin (Sigma). p-RTK arrays were performed according to the manufacturer's recommendations (Human Phospho-RTK Array Kit, R&D Systems). For PDGFR $\beta$  immunohistochemistry, paraffin-embedded formalin-fixed tissue sections were antigen-retrieved, incubated with a PDGFR $\beta$  antibody followed by horseradish peroxidase-conjugated secondary antibody (Envision System, DakoCytomation). Immunocomplexes were visualized using the DAB (3,3'-diaminobenzidine) peroxidase method and nuclei haematoxylin-counterstained. For activated RAS pull-down, lysates were incubated with beads coupled to glutathione-S-transferase (GST)-RAF-1-RAS-binding domain of RAF1 (RBD) (Thermo) for 1 h at 4 °C.

**RNA quantifications.** For real-time quantitative PCR, total RNA was extracted and cDNA quantified. Data were normalized to tubulin and *GAPDH* levels. Relative expression is calculated using the delta-Ct method. For RNA expression profiling, total RNAs were extracted, and generated cDNAs were fragmented, labelled and hybridized to the GeneChip Human Gene 1.0 ST Arrays (Affymetrix). Expression data were normalized, background-corrected, and log<sub>2</sub>-transformed for parametric analysis. Differentially expressed genes were identified using significance analysis of microarrays (SAM) with the R package 'samr' (false discovery rate (FDR) < 0.05; fold change > 2).

**Cell cycle and apoptosis.** For cell cycle analysis, cells were fixed, permeabilized and stained with propidium iodide (BD Pharmingen). Cell cycle distribution was analysed by Cell Quest Pro and Modifit software. For apoptosis, cells were co-stained with Annexin V-V450 and propidium iodide (BD Pharmingen). Data were analysed with the FACS Express V2 software.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 27 July; accepted 2 November 2010.

Published online 24 November 2010.

1. Davies, H. *et al.* Mutations of the *BRAF* gene in human cancer. *Nature* **417**, 949–954 (2002).

2. Flaherty, K. T. *et al.* Inhibition of mutated, activated BRAF in metastatic melanoma. *N. Engl. J. Med.* **363**, 809–819 (2010).
3. Jänne, P. A., Gray, N. & Settleman, J. Factors underlying sensitivity of cancers to small-molecule kinase inhibitors. *Nature Rev. Drug Discov.* **8**, 709–723 (2009).
4. Montagut, C. *et al.* Elevated CRAF as a potential mechanism of acquired resistance to BRAF inhibition in melanoma. *Cancer Res.* **68**, 4853–4861 (2008).
5. Poulikakos, P. I., Zhang, C., Bollag, G., Shokat, K. M. & Rosen, N. RAF inhibitors transactivate RAF dimers and ERK signalling in cells with wild-type BRAF. *Nature* **464**, 427–430 (2010).
6. Søndergaard, J. N. *et al.* Differential sensitivity of melanoma cell lines with *BRAF*<sup>V600E</sup> mutation to the specific Raf inhibitor PLX4032. *J. Transl. Med.* **8**, 39–50 (2010).
7. Halaban, R. *et al.* PLX4032, a selective *BRAF*<sup>V600E</sup> kinase inhibitor, activates the ERK pathway and enhances cell migration and proliferation of *BRAF*<sup>WT</sup> melanoma cells. *Pigment Cell Melanoma Res.* **23**, 190–200 (2010).
8. Hatzivassiliou, G. *et al.* RAF inhibitors prime wild-type RAF to activate the MAPK pathway and enhance growth. *Nature* **464**, 431–435 (2010).
9. Heidorn, S. J. *et al.* Kinase-dead BRAF and oncogenic RAS cooperate to drive tumor progression through CRAF. *Cell* **140**, 209–221 (2010).
10. Bollag, G. *et al.* Clinical efficacy of a RAF inhibitor needs broad target blockade in *BRAF*-mutant melanoma. *Nature* **467**, 596–599 (2010).
11. Tsai, J. *et al.* Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity. *Proc. Natl Acad. Sci. USA* **105**, 3041–3046 (2008).
12. Whittaker, S. *et al.* Gatekeeper mutations mediate resistance to BRAF-targeted therapies. *Sci. Transl. Med.* **2**, 35ra41 (2010).
13. Pratilas, C. A. *et al.* *BRAF*<sup>V600E</sup> is associated with disabled feedback inhibition of RAF-MEK signaling and elevated transcriptional output of the pathway. *Proc. Natl Acad. Sci. USA* **106**, 4519–4524 (2009).
14. Packer, L. M., East, P., Reis-Filho, J. S. & Marais, R. Identification of direct transcriptional targets of BRAF/MEK signalling in melanoma. *Pigment Cell Melanoma Res.* **22**, 785–798 (2009).
15. Wu, E. *et al.* Comprehensive dissection of PDGF-PDGFR signaling pathways in PDGFR genetically defined cells. *PLoS ONE* **3**, e3794 (2008).
16. Smith, G. *et al.* Activating K-Ras mutations without 'hotspot' codons in sporadic colorectal tumours - implications for personalised cancer medicine. *Br. J. Cancer* **102**, 693–703 (2010).
17. Emery, C. M. *et al.* MEK1 mutations confer resistance to MEK and B-RAF inhibition. *Proc. Natl Acad. Sci. USA* **106**, 20411–20416 (2009).
18. Dumaz, N. *et al.* In melanoma, RAS mutations are accompanied by switching signaling from BRAF to CRAF and disrupted cyclic AMP signaling. *Cancer Res.* **66**, 9483–9491 (2006).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature). A figure summarizing the main result of this paper is also included as SI.

**Acknowledgements** We are grateful to G. Bollag and P. Lin (Plexxikon) for providing PLX4032, J. S. Economou for biopsies, B. Comin-Anduix for FACS assistance, S. Mok for assistance with virus production, C. Ng for tissue acquisition and culture establishment, R. Huang for patient tissue processing, N. Doan for immunohistochemistry, P. Mischel for discussion, B. Chmielowski for coordinated patient care, T.L. Toy for technical help with library generation for deep sequencing, and B. Harry for help with analysis of deep sequence data. R.S.L. acknowledges funding from the following: Dermatology Foundation, Burroughs Wellcome Fund, STOP CANCER Foundation, Margaret E. Early Medical Trust, Ian Copeland Memorial Melanoma Fund, V Foundation for Cancer Research, Melanoma Research Foundation, American Skin Association, Caltech-UCLA Joint Center for Translational Medicine, Wesley Coyle Memorial Fund, and Melanoma Research Alliance. R.N. is supported by a post-doctoral fellowship from the T32 Tumor Immunology Training Grant (S. Dubinett). A.R. is supported by the California Institute for Regenerative Medicine (CIRM), the Jonsson Cancer Center Foundation (JCCF), and Caltech-UCLA Joint Center for Translational Medicine. Array and sequence work were performed within the Jonsson Comprehensive Cancer Center Gene Expression Shared Resource. Patient-informed consent was obtained for the research performed in this study. We would like to thank all the patients that participated in this study.

**Author Contributions** R.N., H.S., Q.W., X.K., H.L., Z.C. designed and performed experiments and analysed data. M.L. helped analyse data. R.C.K., N.A., H.S., T.C., G.M., J.A.S., and A.R. provided reagents. S.F.N. helped design experiments and interpreted data. A.R. and R.S.L. designed research aims. H.L., S.F.N., G.M., J.A.S. and A.R. helped write the paper. R.S.L. designed and performed experiments, analysed data, provided reagents, and wrote the paper.

**Author Information** Gene expression and copy number data are deposited at Gene Expression Omnibus under accession numbers GSE24862 and GSE24890, respectively. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details accompany the full-text HTML version of the paper at [www.nature.com/nature](http://www.nature.com/nature). Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to R.S.L. ([rio@mednet.ucla.edu](mailto:rio@mednet.ucla.edu)).



## METHODS

**Cell culture, lentiviral constructs and infections.** All cell lines were maintained in DMEM with 10% or 20% (short-term cultures) heat-inactivated FBS (Omega Scientific) and  $2 \text{ mmol l}^{-1}$  glutamine in humidified, 5%  $\text{CO}_2$  incubator. To derive PLX4032-resistant sub-lines, M229 and M238 were seeded at low cell density and treated with PLX4032 at  $1 \mu\text{M}$  every 3 days for 4–6 weeks and clonal colonies were then isolated by cylinders. M249 R was derived by successive titration of PLX4032 up to  $10 \mu\text{M}$ . PLX4032-resistant sub-lines and short-term cultures were replenished with  $1 \mu\text{M}$  PLX4032 every 2 to 3 days. shRNAs were sub-cloned into the lentiviral vector pLL3.7. N-RAS(Q61K) mutant overexpression construct was made by PCR-amplifying from M249 R4 cDNA and sub-cloning into the lentiviral vector (UCLA Vector Core), creating pRRLsin.cPPT.CMV.hTERT.IRES.GFP-Flag-Q61K-NRAS. Wild-type PDGFR $\beta$  overexpression construct was PCR-amplified from cDNA and sub-cloned into a lentiviral vector (Clontech), creating pLVX-Tight-Puro-PDGFR $\beta$ -Myc. Lentiviral constructs were co-transfected with three packaging plasmids into HEK293T cells. Infections were carried out with protamine sulphate.

**Cellular proliferation, drug treatments and siRNA transfections.** Cell proliferation experiments were performed in a 96-well format (five replicates), and baseline quantification performed at 24 h after cell seeding along with initiation of drug treatments (72 h). Stocks and dilutions of PLX4032 (Plexxikon), AZD6244 (Selleck Chemicals) and U0126 (Promega) were made in DMSO. siRNA pool (Dharmacon) transfections were carried out in 384-well format. TransIT transfection reagent (Mirus) was added to each well and incubated at  $37^\circ\text{C}$  for 20 min. Subsequently, cells were reverse transfected, and the mixture was incubated for 51–61 h at  $37^\circ\text{C}$ . Cells were quantified using CellTiter 96 Aqueous One Solution (Promega) or CellTiter-GLO Luminescence (Promega) following the manufacturer's recommendations.

**Protein detection.** Cell lysates for western blotting were made in RIPA (Sigma) with protease inhibitor cocktail (Roche) and phosphatase inhibitor cocktails I and II (Santa Cruz Biotechnology). Western blots were probed with antibodies against p-MEK1/2 (S217/221), total MEK1/2, p-ERK1/2 (T202/Y204), total ERK1/2, PDGFR $\beta$ , and EGFR (all from Cell Signaling Technologies), B-RAF and N-RAS (Santa Cruz Biotechnology), pan-RAS (Thermo Scientific) and tubulin (Sigma). p-RTK arrays were performed according to the manufacturer's recommendations (Human Phospho-RTK Array Kit, R&D Systems). For PDGFR $\beta$  immunohistochemistry, paraffin-embedded formalin fixed tissue sections were subjected to antigen retrieval and incubated with a rabbit monoclonal anti-PDGFR $\beta$  antibody (Cell Signaling Technology) followed by labelled anti-rabbit polymer horseradish peroxidase (Envision System, Dako Cytomation). Immunocomplexes were visualized using the DAB (3,3'-diaminobenzidine) peroxidase method and nuclei haematoxylin-counterstained.

**In vitro kinase assay.** Cells were harvested and protein lysates prepared in a NP40-based buffer before subjected to immunoprecipitation (IP). IP beads were then resuspended in ADBI buffer (with Mg/ATP cocktail) and incubated with an inactive, recombinant MEK1 or a truncated RAF-1 (positive control) (Millipore), and with DMSO or  $1 \mu\text{M}$  PLX4032 for 30 min at  $30^\circ\text{C}$ . The beads were subsequently pelleted and the supernatant resuspended in sample buffer for western blotting to detect p-MEK and total MEK.

**Activated RAS pull-down assay.** Melanoma lysates were incubated with glutathione agarose beads coupled to  $80 \mu\text{g}$  GST-RAF-1-RBD (Thermo) for 1 h at  $4^\circ\text{C}$ . As controls, Pt48 R lysate was pre-incubated with either  $0.1 \text{ mM}$  GTP $\gamma\text{S}$  (positive control) or  $1 \text{ mM}$  GDP (negative control) in the presence of  $10 \text{ mM}$  EDTA (pH 8.0) at  $30^\circ\text{C}$  for 15 min. Reactions were terminated by adding  $60 \text{ mM}$   $\text{MgCl}_2$ . After washing with Wash Buffer (Thermo), proteins bound to beads were eluted by protein sample buffer. RAS or NRAS levels were detected by immunoblotting.

**Quantitative real-time PCR for relative RNA levels.** Total RNA was extracted using the RiboPure Kit (Ambion), and reverse transcription reactions were performed using the SuperScript First-Strand Synthesis System (Invitrogen). Real-time PCR analyses were performed using the iCycler iQ Real Time PCR Detection System (BioRad) (Supplementary Table 5). To discriminate specific from non-specific cDNA products, a melting curve was obtained at the end of each run. Data were normalized to tubulin and/or *GAPDH* levels in the samples in duplicates. Relative expression is calculated using the delta-Ct method using the following equations:  $\Delta\text{Ct}(\text{Sample}) = \text{Ct}(\text{Target}) - \text{Ct}(\text{Reference})$ ; relative quantity =  $2^{-\Delta\text{Ct}}$ .

**Quantitative real-time PCR for relative DNA copy numbers.** gDNAs were extracted using the FlexiGene DNA Kit (Qiagen) (Human Genomic DNA-Female, Promega). NRAS relative copy number was determined by quantitative PCR (cycle conditions available upon request) using the MyiQ single colour Real-Time PCR Detection System (Bio-Rad). Total DNA content was estimated by assaying  $\beta$ -globin for each sample (Supplementary Table 5), and  $20 \text{ ng}$  of gDNA was mixed with the SYBR Green QPCR Master Mix (Bio-Rad) and  $2 \text{ pmol l}^{-1}$  of each primer.

**Sequencing.** gDNAs were isolated using the Flexi Gene DNA Kit (QIAGEN) or the QIAamp DNA FFPE Tissue Kit. B-RAF and RAS genes were amplified from genomic DNA by PCR. PCR products were purified using QIAquick PCR Purification Kit (QIAGEN) followed by bi-directional sequencing using BigDye v1.1 (Applied Biosystems) in combination with a 3730 DNA Analyzer (Applied Biosystems). PDGFR $\beta$  was amplified from cDNA by PCR and sequenced (primers listed in Supplementary Table 1).

**B-RAF ultra-deep sequencing.** Exon-based amplicons were generated using Platinum high-fidelity *Taq* polymerase, and libraries were prepared following the Illumina library generation protocol version 2.3. For each sample, one library was generated with 18 exons pooled at equal molarity and another library was generated for exon 13 only for validation purpose. Each library was indexed with an unique four base long barcode within the custom made Illumina adaptor. All 10 indexed samples were pooled and sequenced on one lane of Illumina GAIIx flow-cell for single-end 76 base pairs. For error rate estimation, phiX174 genome was spiked in. Base-calling was performed by Illumina RTA version 1.8.70. Alignment was performed using the Novocraft Short Read Alignment Package version 2.06 (<http://www.novocraft.com/index.html>). First, all reads were aligned to the phiX174 reference genome downloaded from the NCBI. The mismatch rates at each position of the reads were calculated to estimate the error rate of the sequencer (set at 1.67% or five standard deviations, SD) based on the phiX genome data (mean error rate = 0.57%, s.d. = 0.22%). Then, the .qseq.txt files were converted into .fastq file using a custom script (available on request) and during this process, the first 5 bases (unique 4-base barcode and the T at the fifth position) were stripped off from the reads and concatenated to the read name. The .fastq file was parsed into 10 .fastq files for each barcode and only the reads with the first 5 bases perfectly matching any of the 10 barcodes were included. Each .fastq file was aligned to chromosome 7 fasta file, generated from the Human Genome reference sequence (hg18, March 2006, build 36.1) downloaded from the Broad Institute ([ftp://ftp.broadinstitute.org/pub/gsa/gatk\\_resources.tgz](ftp://ftp.broadinstitute.org/pub/gsa/gatk_resources.tgz)) using the Novoalign program. Base calibration option was used, and the output format was set to SAM. Using SAMtools (<http://samtools.sourceforge.net/>), the .sam files of each lane were converted to .bam files and sorted, followed by removal of potential PCR duplicates using Picard (<http://picard.sourceforge.net/>). The true background rate was inferred from analysis of independent exon 13 amplicons. None of the 14 positions within exon 13 that had non-reference allele frequency (NAF) > 1.67% in all-exon-samples were validated in the exon13-only samples and vice versa for the one position in the exon 13-only sample, inferring that the true background error rate could be higher at 4.81% (5s.d., mean error rate = 2.72%, s.d. = 0.4%). In total, 12 positions had NAF > 4.81%, and none of them occurred at the same position. We note that the four sample gDNAs extracted from formalin-fixed paraffin-embedded (FFPE) blocks had 5–6 times more variants with NAF above background than the sample extracted from frozen tissue, and the 12 positions with NAF > 4.81% were scattered only across the FFPE samples. The numbers of variants within and outside the kinase domain were not significantly different.

**B-RAF deep sequence from whole exome sequence analysis.** Genomic libraries were generated following the Agilent SureSelect Human All Exon Kit Illumina Paired-End Sequencing Library Prep Version 1.0.1 protocol at the UCLA Genome Center. Agilent SureSelect All Exon ICGC version was used for capturing ~50 megabase (Mb) exome. The Genome Analyzer IIX (GAIIx) was run using standard manufacturer's recommended protocols. Base-calling was done by Illumina RTA version 1.6.47. Two lanes of Illumina single end (SE) run were generated for each of Pt111-001 normal, baseline and DP2 samples, and one lane of Illumina paired end (PE) run was generated for each of Pt111-001 DP1, DP3 as well as Pt111-010 normal, baseline, DP1 and DP2 samples. Alignment was performed using the Novocraft Short Read Alignment Package version 2.06. Human Genome reference sequence (hg18, March 2006, build 36.1), downloaded from the UCSC genome database located at <http://genome.ucsc.edu> and mirrored locally, was indexed using novoindex program (-k 14 -s 3). Novoalign program was used to align each lane's qseq.txt file to the reference genome. Base calibration option and adaptor stripping option for paired-end run were used and the output format was set to SAM. Using SAMtools (<http://samtools.sourceforge.net/>), the .sam files of each lane were converted to .bam files, sorted and merged for each sample and potential PCR duplicates were removed using Picard (<http://picard.sourceforge.net/>). The .bam files were filtered for SNV calling and small INDEL calling to reduce the likelihood of using spuriously mis-mapped reads to call the variants. For the .bam file to call SNVs, the last 5 bases were trimmed and only the reads lacking indels were retained. For the .bam file to call small INDELs, only the reads containing one contiguous INDEL but not positioned at the beginning or the end of the read were retained. SOAP consensus-calling model implemented in SAMtools was used to call the variants, both SNVs and indels, and generate the .pileup files for each .bam file. Coding regions  $\pm 2 \text{ bp}$  of B-RAF gene were extracted



from the .pileup files and the reads were manually examined for rare variants (non reference alleles).

**Microarray data generation and analysis.** Total RNAs were extracted using the RiboPure Kit (Ambion) from cells (DMSO or PLX4032, 1  $\mu$ M, 6 h). cDNAs were generated, fragmented, biotinylated, and hybridized to the GeneChip Human Gene 1.0 ST Arrays (Affymetrix). The arrays were washed and stained on a GeneChip Fluidics Station 450 (Affymetrix); scanning was carried out with the GeneChip Scanner 3000 7G; and image analysis with the Affymetrix GeneChip Command Console Scan Control. Expression data were normalized, background-corrected, and summarized using the RMA algorithm implemented in the Affymetrix Expression ConsoleTM version 1.1. Data were log-transformed (base 2) for parametric analysis. Clustering was performed with MeV 4.4, using unsupervised hierarchical clustering analysis on the basis of Pearson correlation and complete/average linkage clustering. Differentially expressed genes were identified using significance analysis of microarrays (SAM) with the R package 'samr' (R 2.9.0; FDR < 0.05; fold change greater than 2). To identify and rank pathways enriched among differentially expressed genes, *P*-values (Fisher's exact test) were calculated for gene sets with at least 20% differentially expressed genes. Curated gene sets of canonical pathways in the Molecular Signatures Database (MSigDB) were used.

**Copy number variation analysis.** Illumina HumanExon510S-DUO bead arrays (Illumina) were performed following the manufacturer's protocol. Scanned array data were imported into BeadStudio software (Illumina), where signal intensities for samples were normalized against those for reference genotypes. Log<sub>2</sub> ratios were calculated, and data smoothed using the median with window size of 10 and step size of five probes.

**Cell cycle and apoptosis analysis.** All infected cells were replenished with PLX4032 24 h after infections (M229 R5 treated with AZD6244 to inhibit rebound p-ERK on PDGFR $\beta$  KD), fixed, permeabilized, and treated with RNase (Qiagen). Cells were stained with 50 mg ml<sup>-1</sup> propidium iodide (BD Pharmingen) and the distribution of cell cycle phases was determined by Cell Quest Pro and ModiFit software. For apoptosis, post-infection cells were stained with Annexin V-V450 (BD Pharmingen) and propidium iodide for 15 min at room temperature. Flow cytometry data were analysed by the FACS Express V2 software.

**Image acquisition and data processing.** Statistical analyses were performed using InStat 3 Version 3.0b (GraphPad Software), and graphical representations using DeltaGraph or Prism (Red Rock Software). An Optronics camera system was used in conjunction with Image-Pro Plus software (MediaCybernetics) and Adobe Photoshop 7.0.

# Crystal structure of bacterial RNA polymerase bound with a transcription inhibitor protein

Shunsuke Tagami<sup>1,2</sup>, Shun-ichi Sekine<sup>1,2,3</sup>, Thirumananseri Kumarevel<sup>4†</sup>, Nobumasa Hino<sup>2</sup>, Yuko Murayama<sup>1,2</sup>, Syunsuke Kamegamori<sup>1,2</sup>, Masaki Yamamoto<sup>5</sup>, Kensaku Sakamoto<sup>2</sup> & Shigeyuki Yokoyama<sup>1,2,3</sup>

The multi-subunit DNA-dependent RNA polymerase (RNAP) is the principal enzyme of transcription for gene expression. Transcription is regulated by various transcription factors. Gre factor homologue 1 (Gfh1), found in the *Thermus* genus, is a close homologue of the well-conserved bacterial transcription factor GreA, and inhibits transcription initiation and elongation by binding directly to RNAP<sup>1–8</sup>. The structural basis of transcription inhibition by Gfh1 has remained elusive, although the crystal structures of RNAP and Gfh1 have been determined separately<sup>6–9</sup>. Here we report the crystal structure of *Thermus thermophilus* RNAP complexed with Gfh1. The amino-terminal coiled-coil domain of Gfh1 fully occludes the channel formed between the two central modules of RNAP; this channel would normally be used for nucleotide triphosphate (NTP) entry into the catalytic site. Furthermore, the tip of the coiled-coil domain occupies the NTP  $\beta$ - $\gamma$  phosphate-binding site. The NTP-entry channel is expanded, because the central modules are ‘ratcheted’ relative to each other by  $\sim 7^\circ$ , as compared with the previously reported elongation complexes. This ‘ratcheted state’ is an alternative structural state, defined by a newly acquired contact between the central modules. Therefore, the shape of Gfh1 is appropriate to maintain RNAP in the ratcheted state. Simultaneously, the ratcheting expands the nucleic-acid-binding channel, and kinks the bridge helix, which connects the central modules. Taken together, the present results reveal that Gfh1 inhibits transcription by preventing NTP binding and freezing RNAP in the alternative structural state. The ratcheted state might also be associated with other aspects of transcription, such as RNAP translocation and transcription termination.

RNAP synthesizes RNA complementary to the template DNA (Supplementary Fig. 1a). Crystallographic studies of RNAPs from thermophilic bacteria and RNAP II (Pol II) from the yeast *Saccharomyces cerevisiae* have revealed the overall structure of RNAP, which resembles a crab’s claw<sup>9–11</sup> (Supplementary Fig. 1b). In the transcribing RNAP (elongation complex, EC), the nascent RNA strand remains bound to the template DNA strand, forming an 8–9 base pair DNA•RNA hybrid. The DNA•RNA hybrid and the downstream DNA duplex are tightly held in the ‘primary channel’ formed between the pincers of the crab claw, in the EC structures of the *S. cerevisiae* and *T. thermophilus* RNAPs<sup>12–21</sup>. The catalytic site of nucleotide addition resides at the joint of the pincers. The substrate NTP is considered to enter the catalytic site through a pore (the ‘secondary channel’) on the back side of the crab claw (Supplementary Fig. 1b). The two pincers are connected by a long  $\alpha$ -helix (the bridge helix), which is located at the junction of the DNA•RNA hybrid-binding site, the downstream DNA-binding site, and the secondary channel. The bridge helix is inherently flexible, adopting both the continuously-helical and kinked conformations<sup>9–11,22,23</sup>.

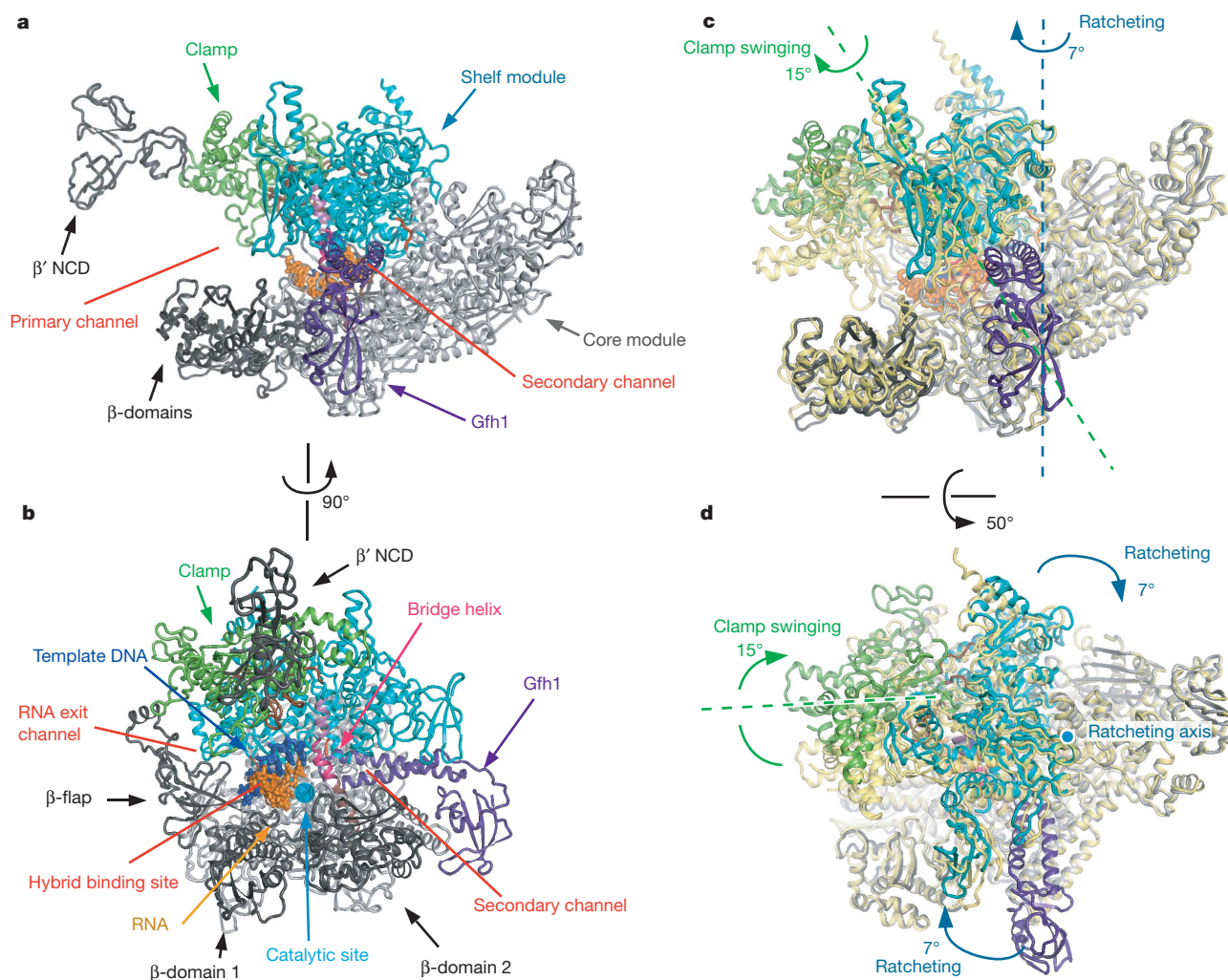
In a previous study<sup>24</sup>, we successfully crystallized *T. thermophilus* RNAP together with DNA, RNA and Gfh1, and collected X-ray

diffraction data sets for two  $P2_1$  crystals (crystals 1 and 2). The nucleic-acid scaffolds employed for the crystallization included the downstream duplex DNA, the DNA•RNA hybrid, and an upstream RNA hairpin 10 or 11 nucleotides (nt) from the RNA 3’ end (Supplementary Text 1, Supplementary Fig. 2a)<sup>24</sup>. In the present study, we determined the structures of the quaternary complex of RNAP•DNA•RNA•Gfh1 (EC•Gfh1) (Fig. 1a, b, Supplementary Table 1, and Methods). The structures of the three independent RNAP molecules in the asymmetric units of crystals 1 and 2 are all similar to each other (Supplementary Text 2, Supplementary Figs 3–5). RNAP and Gfh1 showed clear electron densities (Supplementary Fig. 6), and thus the inhibition mechanisms of Gfh1 were unambiguously revealed, as described below. In contrast, the electron densities of both the DNA and RNA were weak, so we only built the partial models (Supplementary Text 3, Supplementary Figs 2, 7).

The *S. cerevisiae* Pol II structure consists of four rigid modules, ‘core’, ‘shelf’, ‘clamp’ and ‘jaw-lobe’, which are mobile relative to each other<sup>10</sup>. The rigid modules of the bacterial RNAP were defined previously, on the basis of the structures of the *Thermus aquaticus* core enzyme and the *T. thermophilus* holoenzyme<sup>11,25,26</sup> (Supplementary Text 4). However, the conformations of the *T. thermophilus* RNAP in the present EC•Gfh1 structures differ appreciably from those in the previously reported structures of the core enzyme, the holoenzyme and ECs<sup>13,14</sup> (Fig. 1c, d). The large conformational differences enabled us to redefine the rigid-body modules of *T. thermophilus* RNAP (Fig. 1a, b, Supplementary Table 2, Supplementary Text 4, Supplementary Figs 8, 9). These rigid modules include the ‘core’, ‘shelf’, and ‘clamp’ modules, which generally correspond well to those in Pol II. The exceptions are that the active-site and dock domains belong to the ‘shelf’ module, rather than to the ‘core’ module, and the  $\beta$ -domain 1 (or the protrusion domain in Pol II) and the  $\beta$  flap domain (or the wall domain in Pol II) are not included in the ‘core’ module in *T. thermophilus* RNAP (Supplementary Text 4, Supplementary Fig. 8). The core and shelf modules are the main structural elements that form the primary, secondary and RNA-exit channels (Fig. 1a, b, Supplementary Text 4). Gfh1 is accommodated in the secondary channel and its exterior region (Fig. 1a, b, Supplementary Fig. 6). The N-terminal domain (NTD) of Gfh1 forms a coiled coil and is inserted into the secondary channel. The major relative movements among the rigid-body modules are the ‘ratcheting’ between the shelf and core modules and the ‘swinging’ of the clamp relative to the shelf module (Fig. 1c, d, Supplementary Movies 1, 2), as described below in more detail.

The clamp module is connected to the shelf module by four loops (switches 1, 2, 4 and 5, Supplementary Table 2). The protruded clamp module swings relative to the shelf module by about  $15^\circ$  around the centre near switch 5, and is further tilted by about  $5^\circ$  in EC•Gfh1 (Fig. 1a, b, Supplementary Fig. 10, Supplementary Movie 2). Consequently, the region of the primary channel outside the hybrid-binding site is

<sup>1</sup>Department of Biophysics and Biochemistry, Graduate School of Science, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. <sup>2</sup>RIKEN Systems and Structural Biology Center, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan. <sup>3</sup>Laboratory of Structural Biology, Graduate School of Science, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. <sup>4</sup>Structural and Molecular Biology Laboratory, RIKEN SPring-8 Center, Harima Institute, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan. <sup>5</sup>SR Life Science Instrumentation Unit, RIKEN SPring-8 Center, Harima Institute, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan. <sup>†</sup>Present address: Biometal Science Laboratory, RIKEN SPring-8 Center, Harima Institute, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan.



**Figure 1 | Structure of EC•Gfh1.** **a, b,** Overall structure of *T. thermophilus* EC•Gfh1 in two orientations. **c, d,** Superposition of EC•Gfh1 and EC (PDB 2O5I, yellow). The core modules of the two structures are superposed minimizing the root mean square deviation (RMSD) between C $\alpha$  atoms. Two orientations are shown. The same colour scheme is used in all figures (RNA,

orange; template DNA, blue; Gfh1, purple; shelf module, cyan; core module, grey; clamp, green; switches 1–5, brown; hinge loop, red; other domains, dark grey). The three regions of the bridge helix are coloured differently (N-terminal, dark pink; central, hot pink; C-terminal, violet).

widened (Supplementary Text 5, Supplementary Fig. 11). Considering that Gfh1 binding to RNAP occurs on the opposite side of the clamp module, and that Gfh1 does not directly contact the clamp module, the clamp swinging seems to be related to the hairpin structure in the RNA of the nucleic acid scaffolds (Supplementary Text 5).

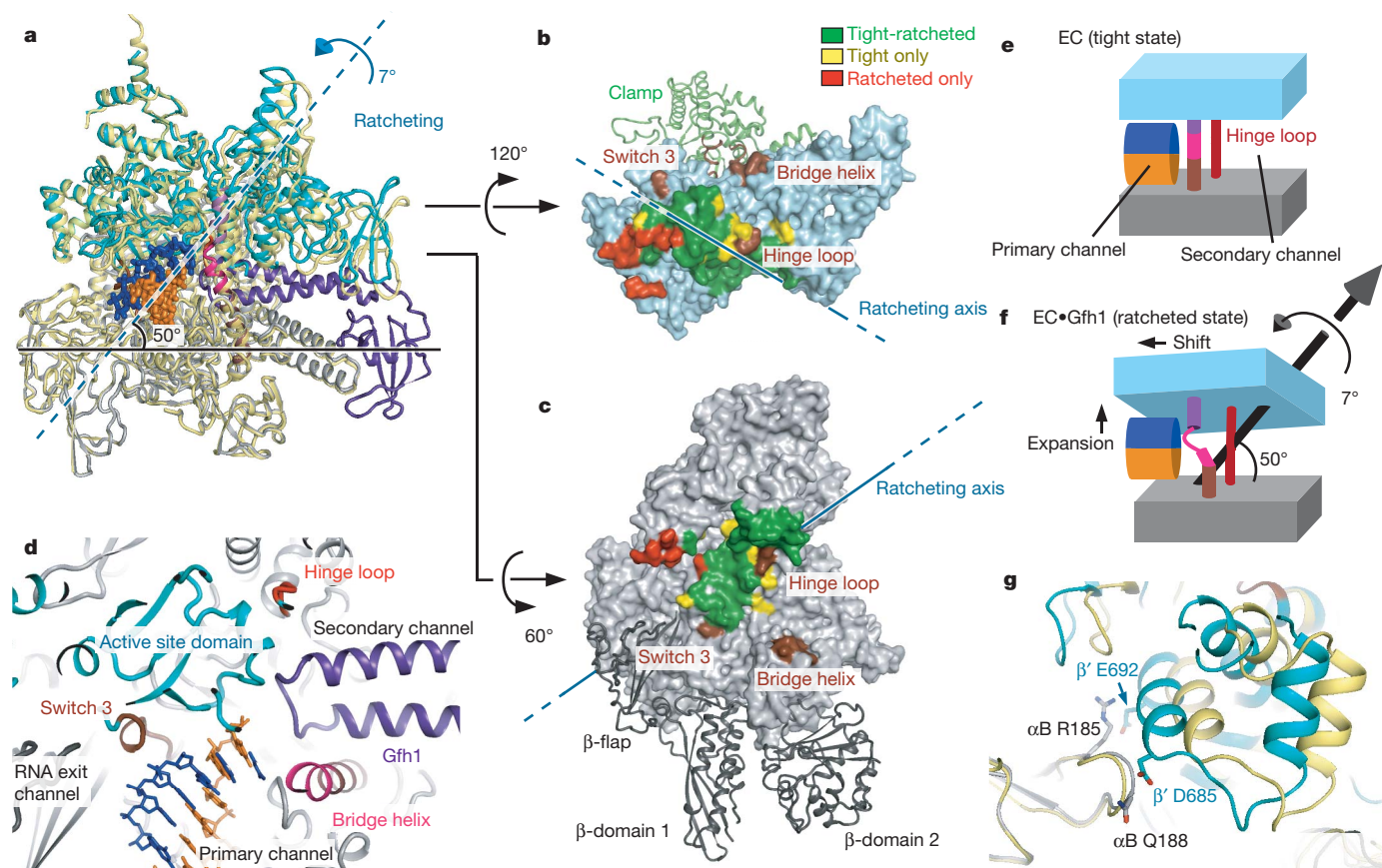
The central body of RNAP is composed of the core and shelf modules, which are rotated by  $\sim 7^\circ$  relative to each other in the present RNAP structures, as compared with the previously-reported EC state<sup>13,14</sup> (Figs 1c, d, 2a, and Supplementary Movie 1). We designate this novel RNAP state as the 'ratcheted' state, and the previous EC state as the 'tight' state. The shelf module is attached to the backboard part of the core module (Fig. 1a), through interfaces of about  $3,800 \text{ \AA}^2$  in the ratcheted state and about  $3,600 \text{ \AA}^2$  in the tight state (Fig. 2b, c). The shelf–core interfaces in the two states are mostly overlapped (shown in green), but there are several contact points specific to either the ratcheted state or the tight state (shown in red and yellow, respectively). The central overlapped area of the interfaces is mainly hydrophobic, and the rotation axis of the ratcheting runs along it (Fig. 2b, c). The ratcheting axis forms an angle of about  $50^\circ$  to the floor of the channels on the core module (Fig. 2a). The core and shelf modules are connected by three peptide segments, the bridge helix, the loop consisting of  $\beta$  Tyr 998–Met 1005 (previously designated as 'switch 3')<sup>10</sup>, and the loop consisting of  $\beta$  Ala 779–Ser 782 (designated hereafter as the 'hinge loop') (Fig. 2d,

Supplementary Fig. 12). As the ratcheting axis runs close to  $\beta'$  Pro 781 (*Escherichia coli*  $\beta'$  Pro 502) within the hinge loop, the conformational change due to the ratcheting is negligibly small around this loop (Fig. 2e, f, Supplementary Fig. 12). By contrast, the bridge helix is much further from the ratcheting axis, as compared with the other two peptide segments (Fig. 2b, c), and the conformational change is significant, as described below in more detail.

On the other hand, the contact points specific to the ratcheted state are also distant from the ratcheting axis (Fig. 2b, c). In particular, an  $\alpha$ -helix of the shelf module ( $\beta'$  685–696) contacts a  $\beta$ -hairpin from one of the  $\alpha$  subunits ( $\alpha$ B 184–191) (Fig. 2g). This contact probably limits the further ratcheting of the shelf module. Therefore, the ratcheted state is mostly at one extremity in the structural spectrum of bacterial RNAP, whereas the tight state observed in previous ECs<sup>13,14</sup> should be the other extremity.

The shelf module ratcheting results in a composite movement that expands the hybrid-binding site and shifts the shelf module forward, relative to the core module (Fig. 2e, f, Supplementary Text 6, Supplementary Fig. 13). The bridge helix exposes its central part ( $\beta'$  1084–1092), but buries its N-terminal part ( $\beta'$  1070–1083) and the carboxy-terminal part ( $\beta'$  1093–1102) within the core and shelf modules, respectively. As the two modules ratchet, the N- and C-terminal parts of the bridge helix shift relative to each other. Consequently, the conformation





**Figure 2 | Ratcheting of the shelf module.** **a**, Superposition of EC•Gfh1 and EC (PDB 2O5I, yellow). The core modules of the two structures are superposed. The clamp,  $\beta'$  NCD, and  $\beta$  domains are not shown. The dashed cyan line indicates the ratcheting axis. **b**, **c**, Interaction surface between the core and the shelf modules; two orientations are shown. **d**, The structures around the active

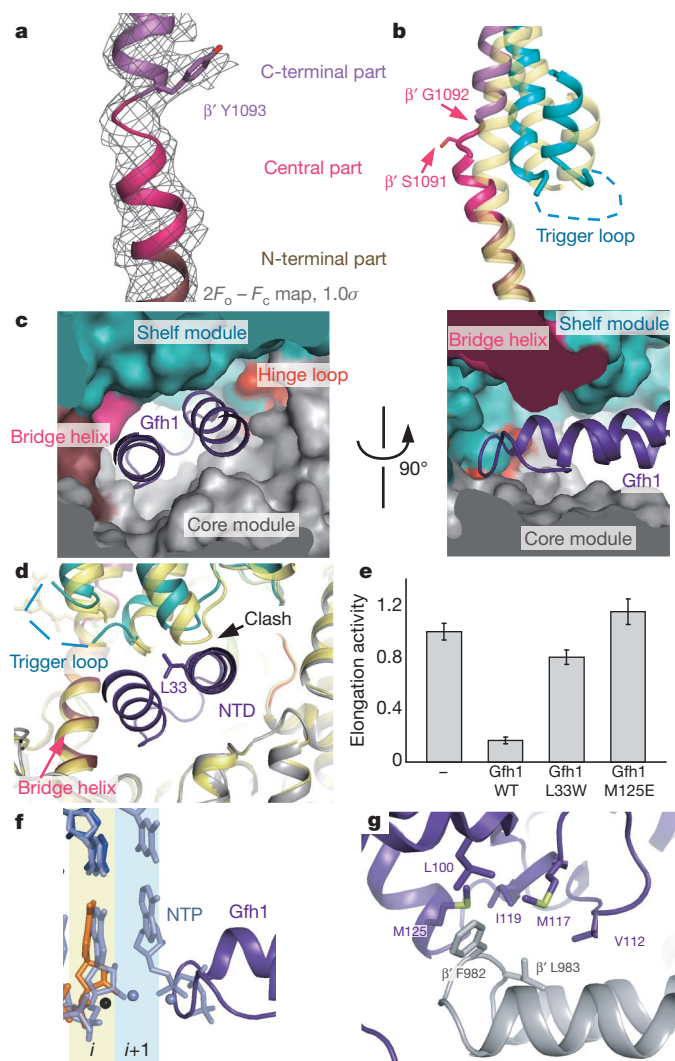
site. The hinge loop is coloured red. **e**, **f**, A schematic drawing representing the structural differences between EC (**e**) and EC•Gfh1 (**f**). **g**, The contacts between the shelf and the core modules specific to the ratcheted state. EC•Gfh1 and EC are superposed by the core modules.

of  $\beta'$  Thr 1088–Gly 1092 in the central part dramatically changes from the continuous  $\alpha$ -helix in the previous ECs, and the two discontinuous  $\alpha$ -helices are connected by the two non-helical residues, Ser 1091–Gly 1092 (Fig. 3a). It is intriguing that mutations of these two residues reportedly affect the RNAP activity<sup>27</sup>. On the other hand, the bridge helix is kinked in the RNAP structures of *T. aquaticus* core enzyme ( $\alpha_2\beta\beta'\omega$ ) and *T. thermophilus* holoenzyme ( $\alpha_2\beta\beta'\omega\sigma$ ) without nucleic acids<sup>9,11</sup>. However, the present kinked conformation is quite different from the previous ones (Supplementary Text 7 and Supplementary Fig. 14). Here, the two  $\alpha$ -helices that are directly connected to the trigger loop are packed against the bridge-helix C-terminal region, and are therefore shifted together with it. To avoid steric hindrance with the tips of these two helices, the residues  $\beta'$  Thr 1088, Ala 1089, Asp 1090 and Ser 1091 (*E. coli*  $\beta'$  Thr 790, Ala 791, Asn 792 and Ser 793) protrude into the DNA•RNA hybrid binding site (Fig. 3b, Supplementary Text 6, Supplementary Fig. 13). Therefore, the conformational change of the bridge helix is likely to occur synchronously with the transition to the ratcheted state (Supplementary Text 7, 8). The direct interaction between Gfh1 and RNAP may affect the fine conformation of the bridge helix, by immobilizing its N-terminal region (see below).

Within the secondary channel, the Gfh1 NTD interacts tightly with parts of the shelf module (including the trigger loop) and the core module (including the secondary-channel coiled coil ( $\beta'$  958–1014) and the N-terminal part of the bridge helix), and fits particularly well with the narrowest region of the secondary channel (Fig. 3c, Supplementary Text 9, Supplementary Fig. 15, Supplementary Movie 1). The interaction between the Gfh1 NTD and the N-terminal part of the bridge helix seems to maintain the straight conformation of the N-terminal part, and to define the kinking point of the bridge helix in EC•Gfh1. It

is impossible for the Gfh1 NTD to bind to the secondary channel in the tight EC in the same manner, as the channel is too narrow (Fig. 3d). Leu 33 of the Gfh1 NTD is located in the narrowest region of the secondary channel. A Gfh1 mutant with Leu 33 replaced by Trp (L33W) lacked transcription inhibition activity, probably because of an inability to bind (Fig. 3e, Supplementary Fig. 16). The bulky side chain of Trp seems to prevent the Gfh1 NTD from penetrating into the channel. This observation also indicates that the secondary channel cannot open beyond the width of the present ratcheted state. Consequently, Gfh1 just fits into the well-defined ratcheted state. Considering that Gfh1 cannot bind to RNAP in the tight state, because of steric hindrance, it is reasonable to postulate that Gfh1 traps a dynamically occurring, ratcheted state of RNAP. To examine this possibility, we performed crosslinking experiments. The results showed that an artificial disulphide bond or photo-crosslink was formed at the ratcheted state-specific interface between the core and shelf modules, even in the absence of Gfh1 (Supplementary Text 10, Supplementary Figs 17–19). Therefore, RNAP might spontaneously and dynamically alter its conformation, from the tight state to the ratcheted state, although the result could also be explained by more localized conformational fluctuations.

Owing to the interaction of the Gfh1 NTD with the narrowest part of the secondary channel, NTP entry would be prevented. Moreover, the presence of the Gfh1 NTD is compatible with the unfolded conformation of the trigger loop, but not with the NTP-induced, folded conformation<sup>14,15</sup> (Supplementary Fig. 20). The tip of the Gfh1 NTD is located near the RNAP active site (Supplementary Text 11, Supplementary Fig. 21). The end of the tip loop occupies the binding site of the  $\beta$ - $\gamma$  phosphate groups in the NTP insertion step of the nucleotide



**Figure 3 | The bridge helix and Gfh1.** **a**, The structure and the electron density of the bridge helix in EC•Gfh1 (crystal 1, RNAP 1). **b**, The bridge helix and the helices neighbouring the trigger loop. *T. thermophilus* EC in the post-translocation state (PDB 2O5I) is superposed, and is coloured yellow. **c**, **d**, The Gfh1 NTD in the secondary channel. **d**, *T. thermophilus* EC in the post-translocation state (PDB 2O5I) is superposed. The black arrow indicates the position where Gfh1 would clash with EC in the tight state. **e**, Inhibition of the nucleotide addition reaction by Gfh1. The nucleotide addition reaction was performed by incubating the transcription elongation complex with 5  $\mu$ M Gfh1 mutants. Error bars, s.d. ( $n \geq 5$ ). **f**, A close-up view of the active site. *T. thermophilus* EC in the NTP-bound state (PDB 2O5I) is superposed, and is coloured light blue.  $i$  and  $i+1$  are the positions of the RNA 3' end and the substrate NTP, respectively, in the nucleotide addition reaction. **g**, Interaction between the tip of the Gfh1 CTD and RNAP.

addition reaction (Fig. 3f). Furthermore, Gfh1 seems to stabilize the kinked bridge helix (Fig. 3b). An antibiotic, streptolydigin, also reportedly inhibited transcription by immobilizing the bridge helix in a fixed conformation<sup>23</sup>. Taken together, these observations provide the explanation for the inhibition of transcription elongation by Gfh1 (Supplementary Text 12, and Supplementary Figs 22, 23).

On the other hand, the C-terminal domain (CTD) of Gfh1 is bound with the edge of the secondary-channel coiled coil of the RNAP (Fig. 1a, b, Supplementary Text 13, Supplementary Fig. 24). The interaction involves the hydrophobic patch on the surface of the Gfh1 CTD (Fig. 3g). Therefore, we prepared a Gfh1 protein with the M125E mutation within the hydrophobic patch, and observed a loss of inhibition activity (Fig. 3e). Therefore, the interaction between the hydrophobic patch of the Gfh1 CTD and the secondary-channel coiled coil is

required for Gfh1 to bind to RNAP. Other Gre factors, such as GreA and GreB, share structural and sequence similarities with Gfh1 (refs 2–8). In particular, the presence of the hydrophobic patch on the CTD is well conserved. In fact, the M124E mutation of *E. coli* GreB (M125E of Gfh1) also reduces the transcript cleavage activity of GreB<sup>3</sup>. Therefore, the Gre factors seem to share a common interaction mode between the hydrophobic patch and the secondary-channel coiled coil, and probably bind to the ratcheted EC in a similar manner to Gfh1 (Supplementary Text 13, 14). Transcript cleavage stimulated by GreA and GreB would be performed in the ratcheted state. Although Pol II EC also changes its structure upon TFIIS binding<sup>18,28</sup>, the observed change is much smaller than the transition to the ratcheted state of *T. thermophilus* EC upon Gfh1 binding (Supplementary Text 15).

The conformational changes of RNAP observed in the present structure, including the shelf module ratcheting and the clamp swinging, might have functional relevance to other stages of the transcription reaction, as the conformational changes should modulate the interactions of RNAP with nucleic acids. Therefore, we suggest that the conformational changes may play distinct roles in RNAP translocation and transcription termination (Supplementary Text 16, 17, Supplementary Figs 25–27). Experimental tests of these hypotheses will be required to assess the importance of these conformational changes in the absence of Gfh1.

## METHODS SUMMARY

**Structure determination.** The structure of crystal 1 was solved by molecular replacement, using the coordinates of RNAP in *T. thermophilus* EC (PDB 2O5I)<sup>13</sup> as the search model. There are three RNAP molecules in the asymmetric unit, and each RNAP is bound with Gfh1. As the relative positions of the Gfh1 NTD and CTD differ from those in free Gfh1 (PDB 2F23)<sup>7</sup>, they were separately placed in the electron density map. The model of the DNA•RNA hybrid was built in the extra electron density in the DNA•RNA hybrid channel. We further remodelled the coordinates of both the proteins and the nucleic acids with the program Coot<sup>29</sup>. Atomic positions and grouped B-factors were refined to 4.1 Å, by using the CNS program<sup>30</sup> (Supplementary Table 1). The refinement converged to  $R$  and  $R_{\text{free}}$  values of 26.2% and 31.8%, respectively; the latter was calculated from randomly-chosen 3% of reflections excluded from the refinement. The structure of crystal 2 was solved by molecular replacement, using the coordinates of the RNAP in crystal 1 as the search model. The models of Gfh1 and the DNA•RNA hybrid were built in the extra electron density. Refinement of the coordinates was performed to 4.3 Å with CNS. The final  $R$  and  $R_{\text{free}}$  values are 31.4% and 33.8%, respectively.

**Transcription inhibition analysis.** We prepared two mutant Gfh1 proteins (L33W and M125E). The elongation complex was reconstituted by incubating *T. thermophilus* RNAP with a nucleic acid scaffold containing template DNA, non-template DNA, and RNA. The RNA was 5'-radiolabelled using T4 polynucleotide kinase and [ $\gamma$ -<sup>32</sup>P]-ATP. The nucleotide addition reaction was performed by incubating the transcription elongation complex in the presence of Gfh1 (the wild type or one of the mutants).

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 16 October 2009; accepted 12 October 2010.**

**Published online 1 December 2010.**

1. Borukhov, S., Polyakov, A., Nikiforov, V. & Goldfarb, A. GreA protein: a transcription elongation factor from *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **89**, 8899–8902 (1992).
2. Stebbins, C. E. *et al.* Crystal structure of the GreA transcript cleavage factor from *Escherichia coli*. *Nature* **373**, 636–640 (1995).
3. Vassilyeva, M. N. *et al.* The carboxy-terminal coiled-coil of the RNA polymerase  $\beta'$ -subunit is the main binding site for Gre factors. *EMBO Rep.* **8**, 1038–1043 (2007).
4. Hogan, B. P., Hartsch, T. & Erie, D. A. Transcript cleavage by *Thermus thermophilus* RNA polymerase. Effects of GreA and anti-GreA factors. *J. Biol. Chem.* **277**, 967–975 (2002).
5. Laptenko, O. & Borukhov, S. Biochemical assays of Gre factors of *Thermus thermophilus*. *Methods Enzymol.* **371**, 219–232 (2003).
6. Lamour, V., Hogan, B. P., Erie, D. A. & Darst, S. A. Crystal structure of *Thermus aquaticus* Gfh1, a Gre-factor paralog that inhibits rather than stimulates transcript cleavage. *J. Mol. Biol.* **356**, 179–188 (2006).
7. Laptenko, O. *et al.* pH-dependent conformational switch activates the inhibitor of transcription elongation. *EMBO J.* **25**, 2131–2141 (2006).

8. Symersky, J. *et al.* Regulation through the RNA polymerase secondary channel. Structural and functional variability of the coiled-coil transcription factors. *J. Biol. Chem.* **281**, 1309–1312 (2006).
9. Zhang, G. *et al.* Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. *Cell* **98**, 811–824 (1999).
10. Cramer, P., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292**, 1863–1876 (2001).
11. Vassylyev, D. G. *et al.* Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* **417**, 712–719 (2002).
12. Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* **292**, 1876–1882 (2001).
13. Vassylyev, D. G., Vassylyeva, M. N., Perederina, A., Tahirou, T. H. & Artsimovitch, I. Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* **448**, 157–162 (2007).
14. Vassylyev, D. G. *et al.* Structural basis for substrate loading in bacterial RNA polymerase. *Nature* **448**, 163–168 (2007).
15. Wang, D., Bushnell, D. A., Westover, K. D., Kaplan, C. D. & Kornberg, R. D. Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* **127**, 941–954 (2006).
16. Westover, K. D., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. *Science* **303**, 1014–1016 (2004).
17. Westover, K. D., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center. *Cell* **119**, 481–489 (2004).
18. Kettenberger, H., Armache, K. J. & Cramer, P. Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIIS. *Mol. Cell* **16**, 955–965 (2004).
19. Brueckner, F. & Cramer, P. Structural basis of transcription inhibition by  $\alpha$ -amanitin and implications for RNA polymerase II translocation. *Nature Struct. Mol. Biol.* **15**, 811–818 (2008).
20. Sydow, J. F. *et al.* Structural basis of transcription: mismatch-specific fidelity mechanisms and paused RNA polymerase II with frayed RNA. *Mol. Cell* **34**, 710–721 (2009).
21. Wang, D. *et al.* Structural basis of transcription: backtracked RNA polymerase II at 3.4 angstrom resolution. *Science* **324**, 1203–1206 (2009).
22. Temiakov, D. *et al.* Structural basis of transcription inhibition by antibiotic streptolydigin. *Mol. Cell* **19**, 655–666 (2005).
23. Tuske, S. *et al.* Inhibition of bacterial RNA polymerase by streptolydigin: stabilization of a straight-bridge-helix active-center conformation. *Cell* **122**, 541–552 (2005).
24. Tagami, S., Sekine, S., Kumarevel, T., Yamamoto, M. & Yokoyama, S. Crystallization and preliminary X-ray crystallographic analysis of *Thermus thermophilus* transcription elongation complex bound to Gfh1. *Acta Crystallogr. F* **66**, 64–68 (2010).
25. Darst, S. A. *et al.* Conformational flexibility of bacterial RNA polymerase. *Proc. Natl Acad. Sci. USA* **99**, 4296–4301 (2002).
26. Murakami, K. S., Masuda, S. & Darst, S. A. Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science* **296**, 1280–1284 (2002).
27. Tan, L., Wiesler, S., Trzaska, D., Carney, H. C. & Weinzierl, R. O. Bridge helix and trigger loop perturbations generate superactive RNA polymerases. *J. Biol.* **7**, 40 (2008).
28. Kettenberger, H., Armache, K. J. & Cramer, P. Architecture of the RNA polymerase II-TFIIIS complex and implications for mRNA cleavage. *Cell* **114**, 347–357 (2003).
29. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
30. Brunger, A. T. Version 1.2 of the Crystallography and NMR system. *Nature Protocols* **2**, 2728–2733 (2007).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This work is based on experiments performed at SPring-8 (with the approval of the Japan Synchrotron Radiation Research Institute) and at the Swiss Light Source (SLS). We thank N. Shimizu for supporting our data collection at SPring-8 beamline BL41XU; T. Tomizaki and C. Schulze-Briese for supporting our data collection at SLS beamline X06SA; and Y. Fujii for assisting with our data collection and for comments. We thank T. Tanaka and K. Sakamoto for assistance in protein preparation. This work was supported in part by a Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Young Scientists (to S.-i.S.), a JSPS Grant-in-Aid for Scientific Research (to S.-i.S. and S.Y.), and the Targeted Proteins Research Program (TPRP), the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan. S.T. was supported by the JSPS Global Centers of Excellence Program (Integrative Life Science Based on the Study of Biosignaling Mechanisms).

**Author Contributions** S.T., S.-i.S., T. K. and S.Y. designed the research. S.T. and S.-i.S. performed the structural analysis. M.Y. supported the structural analysis. S.T., S.-i.S., N.H., S.K. and K.S. performed the disulphide-bonding and/or photo-crosslinking analyses. S.T. and Y.M. performed the biochemical analysis of Gre factors. S.-i.S. created the movies. S.T., S.-i.S. and S.Y. wrote the paper.

**Author Information** The structures of EC•Gfh1 have been deposited in the Protein Data Bank, under accession numbers 3AOH (crystal 1) and 3AOI (crystal 2). Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to S.-i.S. ([sekine@biochem.s.u-tokyo.ac.jp](mailto:sekine@biochem.s.u-tokyo.ac.jp)) or S.Y. ([yokoyama@biochem.s.u-tokyo.ac.jp](mailto:yokoyama@biochem.s.u-tokyo.ac.jp)).



## METHODS

**Structure determination.** Crystallization and data collection were described previously<sup>24</sup>. The data were reprocessed with the XDS program<sup>31</sup>. The structure for crystal 1 was solved by molecular replacement with the program Phaser<sup>32</sup>, using the coordinates of the core enzyme portion of *T. thermophilus* EC (PDB 2O51)<sup>13</sup> as the search model. The asymmetric unit contains three RNAP molecules. Each RNAP was divided into 25–26 rigid bodies, and their positions were refined with the program CNS version 1.2<sup>30,33</sup>. Several of the rigid bodies deviated substantially from the electron density, and they were manually adjusted to the density with the program Coot<sup>29</sup>. For the tip portion of the  $\beta'$  non-conserved domain ( $\beta'$  NCD,  $\beta'$  132–454), the coordinates of the *T. thermophilus* holoenzyme (PDB 3DXJ)<sup>34</sup> were used. Several rounds of rigid body refinement and manual adjustment were performed. In each RNAP molecule, extra electron density, corresponding to Gfh1, was observed. The NTD and CTD coordinates of free Gfh1 (PDB 2F23)<sup>7</sup> were separately placed in the  $2F_o - F_c$  electron density map. One of the RNAP molecules in the asymmetric unit exhibited extra electron density corresponding to the DNA•RNA hybrid in the DNA•RNA hybrid binding site, for which we built the hybrid model. The coordinates of the DNA•RNA hybrid in *S. cerevisiae* EC (PDB 2VUM)<sup>19</sup> were used as the starting model. The electron density for the nucleic acids in the other two RNAP complexes was weak, probably owing to low occupancy and/or high mobility, and therefore, we did not build their models.

The structures of the N- and C-terminal parts of the bridge helix in EC•Gfh1 are similar to those in the previous EC (2O51)<sup>13</sup>, while the central part of the bridge helix in the present complex exhibits a conformational change, due to the ratcheting of the core and shelf modules. The region of  $\beta'$  1086–1090 assumes a helical, but slightly curved conformation, and the model was built by adjusting the corresponding region in the previous EC (2O51) to the electron density. Most parts of the bridge helix ( $\beta'$  1070–1090 and  $\beta'$  1093–1102) maintained the helical conformation.  $\beta'$  Ser 1091 and  $\beta'$  Gly 1092 were placed to link the two discontinuous helices, while fitting their main chains into the electron density. For this rebuilding, the position of  $\beta'$  Tyr 1093, which was identified by the electron density of its large side chain, was helpful (Fig. 3a). The coordinates of both the proteins and the nucleic acids were further refined with the program Coot. The atomic positions and the grouped B-factors were refined to 4.1 Å, by using CNS with strong NCS restraints among the three complexes in the asymmetric unit (Supplementary Table 1). Refinement was monitored by  $R_{\text{free}}$ , calculated from 3% of the reflections that were excluded from the refinement.

The structure for crystal 2 was solved by molecular replacement, using the coordinates of RNAP and Gfh1 in crystal 1 as the search model. A model of the DNA•RNA hybrid was placed in each RNAP in the electron density map, and positional refinement of the coordinates was performed to 4.3 Å with CNS. For the calculation of  $R_{\text{free}}$ , the same reflections as those chosen for crystal 1 were used.

The rigid bodies used in the structural refinement allowed us to define the mobile modules of *T. thermophilus* RNAP. We first superposed the RNAPs in the present EC•Gfh1 and the previous EC (2O51)<sup>13</sup> by certain rigid bodies, and then inspected the rigid bodies that superposed well concurrently. The masses of the co-relocated rigid bodies were defined as modules. Finally, we confirmed that the defined mobile modules relocated separately by the RigidFinder program<sup>35</sup>. **Disulphide-bonding assay for the ratcheted RNAP.** We constructed a plasmid that allows co-expression of the  $\alpha$ ,  $\beta$ ,  $\beta'$  and  $\omega$  subunits of *T. thermophilus* RNAP in *E. coli*, for the preparation of recombinant *T. thermophilus* RNAP (pRpoBCAZ, to be published elsewhere). The recombinant RNAPs (the wild type and the  $\alpha$  Q188C- $\beta'$  D685C mutant) were expressed using this system. They were purified by the procedure used for the natural RNAP core enzyme from *T. thermophilus* cells<sup>36</sup>, except that the cell lysate was heat-treated at 70 °C for 30 min, in order to denature most of the non-thermophilic *E. coli* proteins. Then, the RNAPs were fractionated by polyethyleneimine precipitation, followed by ammonium sulphate precipitation. The recombinant RNAPs were further purified by chromatography on Q-Sepharose and Superdex pg200 columns (GE Healthcare Biosciences).

For the disulphide-bonding analysis, the recombinant RNAP (the wild type or the  $\alpha$  Q188C- $\beta'$  D685C mutant) was dissolved in 75 mM Tris-HCl buffer (pH 8.1), containing 50 mM KCl, 10 mM MgCl<sub>2</sub> and 1 mM DTT. Each of the RNAPs (0.2  $\mu$ M) was incubated with 0.7  $\mu$ M of the nucleic acid scaffold (DNATS/DNANT/RNA14) or 10  $\mu$ M of *T. thermophilus* Gfh1 for 30 min. Then, 2.5 mM glutathione disulphide (GSSG) was added to each RNAP solution for the mild oxidation of Cys residues. The mixtures were analysed by SDS-PAGE, using sample buffer lacking a reducing agent. The formation of a disulphide bond between  $\alpha$  C188 and  $\beta'$  C685 was confirmed by the appearance of an extra band with low mobility, which corresponded to the crosslinked  $\alpha$  and  $\beta'$  subunits (Supplementary Fig. 17). The sequences of the nucleic acids are as follows. The

RNA oligomer: RNA14, UUUUUGAGUCUGCGGCGAU. The DNA oligomers: DNATS, AACATACGGCTCGGACAGAGGTCTGTCTGAATCGATATCGC CGC; DNANT, CGATTACAGACAGGACCTCTGTCCGAGCCGTATGTT. The nucleic acid scaffold was designed by modifying the previously reported EC14 scaffold, which forms a stable elongation complex with *T. thermophilus* RNAP<sup>37</sup>. **Photo-crosslinking assay of the ratcheted RNAP.** *p*-Benzoyl-L-phenylalanine (pBpa) is a photo-crosslinker that can be position-specifically incorporated into a recombinant protein<sup>38,39</sup>. The gene encoding a pBpa-specific variant of *Methanococcus jannaschii* tyrosyl-tRNA synthetase<sup>38</sup>, under the control of the *E. coli* *tyrS* promoter, was cloned in the pACYC184 vector, together with three copies of the *M. jannaschii* amber suppressor tRNA gene<sup>39</sup>, to create a vector for the expression of the pBpa-specific tRNA synthetase and tRNA (ppbpaRS-3MJRI). Each tRNA gene had an *E. coli* *lpp* promoter and an *rrnC* terminator. The artificial operons for overproducing minor tRNA species, including the minor tRNA<sup>Pro</sup>, were described previously<sup>40</sup>, and were cloned in a kanamycin-resistant plasmid carrying the CloDF13-derived replication origin, to create pMINOR2.

The *rpoA* gene, C-terminally tagged with FLAG, was engineered, using a QuikChange mutagenesis kit (Stratagene), to have an amber codon in place of Arg 185, for producing the RNAP  $\alpha$ -subunit with Arg 185 replaced with pBpa (pBpa 185) (ref. 41). The *rpoC* gene was engineered to have a methionine codon in place of Glu 692, for producing the  $\beta'$  subunit with the E692M substitution. The *rpoA* and *rpoC* genes in pRpoBCAZ were replaced by these mutant genes, and the vector was introduced into BL21 Star(DE3) cells (Invitrogen) harbouring the ppbpaRS-3MJRI and pMINOR2 plasmids. The cells were grown in LB medium containing 1 mM pBpa, and the gene expression was induced by the addition of 1 mM IPTG at the mid-log phase. After a further 4-h incubation, the cells were harvested and lysed by sonication in buffer A (40 mM Tris-HCl (pH 7.7), 500 mM NaCl, 10 mM EDTA, 10 mM 2-mercaptoethanol, 5% glycerol, and Complete protease inhibitor cocktail tablets (Roche)). The wild-type and engineered RNAP core enzymes were roughly purified by heat-treatment. For the photo-crosslinking, the proteins were exposed to light at 365 nm for 30 min on ice<sup>41</sup> in a 24-well cell culture plate (BD Biosciences), followed by SDS-PAGE, Coomassie brilliant blue staining, and western blotting with an anti-FLAG antibody (Sigma) (Supplementary Fig. 19a, 19b). The RNAP core enzymes were further purified as described above. Then, 2  $\mu$ M RNAP (wild type or mutant) was incubated with the nucleic acid scaffold (2.5  $\mu$ M) or Gfh1 (10  $\mu$ M) for 30 min at room temperature in 50 mM HEPES-NaOH buffer (pH 7.5), containing 50 mM KCl, 10 mM MgCl<sub>2</sub> and 1 mM DTT, followed by the photo-crosslinking step (Supplementary Fig. 19c).

**Transcription inhibition analysis.** In the previous study, we constructed a plasmid for the expression of wild-type Gfh1 in *E. coli*<sup>24</sup>. The expression plasmids for Gfh1 variants (L33W and M125E) were generated by introducing mutations to the plasmid encoding wild-type Gfh1. In addition, we constructed plasmids for the expression of wild type and mutant *T. thermophilus* GreA in *E. coli*. The wild-type and mutant Gre proteins were expressed as described previously<sup>24</sup>, and were then purified by chromatography on Toyopearl Super-Q and Butyl columns (Tosoh Bioscience). The transcription elongation complex was reconstituted by incubating 0.1  $\mu$ M of *T. thermophilus* RNAP with 0.1  $\mu$ M of the nucleic acid scaffold (DNATS/DNANT/RNA14 or RNA15) for 30 min, where the sequence of RNA15 is UUUUUG AGUCUGCGGCGGAUA. The RNA was 5'-radiolabelled using T4 polynucleotide kinase and [ $\gamma$ -<sup>32</sup>P]-ATP. The nucleotide addition reaction was performed by incubating the transcription elongation complex of RNA14 with 5  $\mu$ M Gfh1 (wild type or a mutant) at 20 °C in 50 mM MES-NaOH buffer (pH 6.5), containing 50 mM KCl, 10 mM MgCl<sub>2</sub>, 1 mM DTT and 20  $\mu$ M ATP and that of RNA15 with 2.5  $\mu$ M Gfh1 (wild type or a mutant) at 55 °C in 50 mM MES-NaOH buffer (pH 6.5), containing 50 mM KCl, 1 mM MgCl<sub>2</sub>, 1 mM DTT and 20  $\mu$ M UTP. The RNA was analysed by denaturing (8 M urea) PAGE.

- Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* **26**, 795–800 (1993).
- McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
- Brünger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
- Mukhopadhyay, J. *et al.* The RNA polymerase “switch region” is a target for inhibitors. *Cell* **135**, 295–307 (2008).
- Abyzov, A., Björnson, R., Felipe, M. & Gerstein, M. RigidFinder: a fast and sensitive method to detect rigid blocks in large macromolecular complexes. *Proteins* **78**, 309–324 (2010).
- Vassilyeva, M. N. *et al.* Purification, crystallization and initial crystallographic analysis of RNA polymerase holoenzyme from *Thermus thermophilus*. *Acta Crystallogr. D* **58**, 1497–1500 (2002).
- Kashkina, E. *et al.* Elongation complexes of *Thermus thermophilus* RNA polymerase that possess distinct translocation conformations. *Nucleic Acids Res.* **34**, 4036–4045 (2006).

38. Chin, J. W., Martin, A. B., King, D. S., Wang, L. & Schultz, P. G. Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **99**, 11020–11024 (2002).
39. Sakamoto, K. *et al.* Genetic encoding of 3-iodo-L-tyrosine in *Escherichia coli* for single-wavelength anomalous dispersion phasing in protein crystallography. *Structure* **17**, 335–344 (2009).
40. Chumpolkulwong, N. *et al.* Translation of 'rare' codons in a cell-free protein synthesis system from *Escherichia coli*. *J. Struct. Funct. Genomics* **7**, 31–36 (2006).
41. Hino, N. *et al.* Protein photo-cross-linking in mammalian cells by site-specific incorporation of a photoreactive amino acid. *Nature Methods* **2**, 201–206 (2005).

# Single-molecule imaging reveals mechanisms of protein disruption by a DNA translocase

Ilya J. Finkelstein<sup>2</sup>, Mari-Liis Visnapuu<sup>2</sup> & Eric C. Greene<sup>1,2</sup>

In physiological settings, nucleic-acid translocases must act on substrates occupied by other proteins, and an increasingly appreciated role of translocases is to catalyse protein displacement from RNA and DNA<sup>1–4</sup>. However, little is known regarding the inevitable collisions that must occur, and the fate of protein obstacles and the mechanisms by which they are evicted from DNA remain unexplored. Here we sought to establish the mechanistic basis for protein displacement from DNA using RecBCD as a model system. Using nanofabricated curtains of DNA and multicolour single-molecule microscopy, we visualized collisions between a model translocase and different DNA-bound proteins in real time. We show that the DNA translocase RecBCD can disrupt core RNA polymerase, holoenzymes, stalled elongation complexes and transcription RNA polymerases in either head-to-head or head-to-tail orientations, as well as EcoRI<sup>E111Q</sup>, *lac* repressor and even nucleosomes. RecBCD did not pause during collisions and often pushed proteins thousands of base pairs before evicting them from DNA. We conclude that RecBCD overwhelms obstacles through direct transduction of chemomechanical force with no need for specific protein–protein interactions, and that proteins can be removed from DNA through active disruption mechanisms that act on a transition state intermediate as they are pushed from one non-specific site to the next.

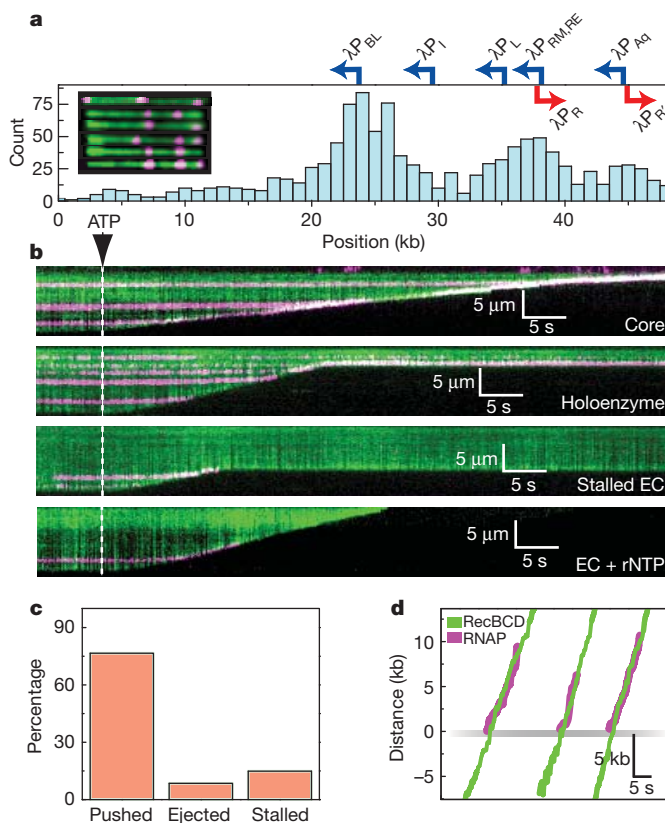
RecBCD is a heterotrimeric translocase involved in initiating homologous recombination and processing stalled replication forks<sup>5,6</sup>. RecB is a 3' → 5' SF1A helicase and contains a nuclease domain for DNA processing, RecD is a 5' → 3' SF1B helicase and RecC holds the complex together and coordinates the response to *cis*-acting Chi (crossover hot-spot instigator) sequences (5'-dGCTGGTGG-3'). RecD is the lead motor before Chi, RecB is the lead motor after Chi and Chi recognition is accompanied by a reduced rate of translocation corresponding to the slower velocity of RecB<sup>7,8</sup>. Chi prompts RecBCD to process DNA, yielding 3' single-stranded DNA overhangs onto which RecA is loaded<sup>7,8</sup>.

We monitored RecBCD activity using total-internal-reflection fluorescence microscopy and a DNA curtain assay that allows us to visualize hundreds of aligned molecules<sup>9</sup> (Supplementary Fig. 1). When assayed on DNA curtains, RecBCD displayed rapid translocation ( $1,484 \pm 167$  base pairs per second ( $\text{bp s}^{-1}$ ), 37 °C, 1 mM ATP,  $N = 100$ ; Supplementary Fig. 1b, c), high processivity ( $36,000 \pm 12,500$  bp) and decreased velocity in response to Chi ( $549 \pm 155 \text{ bp s}^{-1}$ , 37 °C, 1 mM ATP,  $N = 100$ ; Supplementary Fig. 1), in agreement with previous studies<sup>6,7</sup>.

*Escherichia coli* contains ~2,000 molecules of RNA polymerase (RNAP), and  $\geq 65\%$  of these are bound to the bacterial chromosome<sup>10</sup>, making RNAP one of the most commonly encountered obstacles in physiological settings. RNAP is of special interest because it is a high-affinity DNA-binding protein (dissociation constant,  $K_d \approx 10$  pM for  $\lambda P_R$  and 100 pM for  $\lambda P_L$ ) and a powerful translocase capable of moving under an applied load of ~14–25 pN (ref. 11). RNAP survives encounters with replication forks<sup>12–14</sup> and stalls fork progression in head-on collisions<sup>15,16</sup>, suggesting that RNAP is among the most formidable roadblocks encountered *in vivo*. During replication restart, RecBCD translocates towards *oriC*; therefore, most collisions with

RNAP will occur in a head-on orientation, suggesting that to survive these encounters RecBCD would need to exert more force than a replisome.

We used quantum dots (QDs) to fluorescently label RNAP (Supplementary Information). The binding distribution of QD–RNAP holoenzyme overlapped with known promoters (Fig. 1a), promoter targeting was  $\sigma^{70}$  dependent and promoter-bound holoenzymes were highly stable ( $t_{1/2} = 23.2 \pm 1.42$  min (half-life),  $N = 58$ ; Supplementary Fig. 3a, b, c). Core QD–RNAP dissociated when challenged with heparin ( $t_{1/2} = 3.4 \pm 0.03$  s,  $N = 150$ ), whereas promoter-bound



**Figure 1 | RecBCD removes RNAP from DNA.** **a**, Distribution of QD–RNAP bound to  $\lambda$  DNA. Locations of promoters are indicated; those facing left are shown in blue, those facing right are shown in red. The inset shows examples of YOYO-1-stained  $\lambda$  DNA (green) bound by RNAP (magenta). The tethered end of the DNA is on the left, and the free end of the DNA is on the right. kb, kilobase. **b**, Kymograms of RecBCD colliding with RNAP core, holoenzyme, stalled elongation complex (EC), and stalled elongation complex chased with ribonucleoside triphosphate (rNTP). Gaps in magenta traces correspond to quantum dot blinking. In these and all subsequent kymograms, the tethered end of the DNA is at the top, the free end is at the bottom, and buffer flow is from top to bottom. **c**, Distribution of event types. **d**, Tracking data for collisions, with traces aligned at the collisions.

<sup>1</sup>Howard Hughes Medical Institute, Columbia University, New York, New York 10032, USA. <sup>2</sup>Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032, USA.

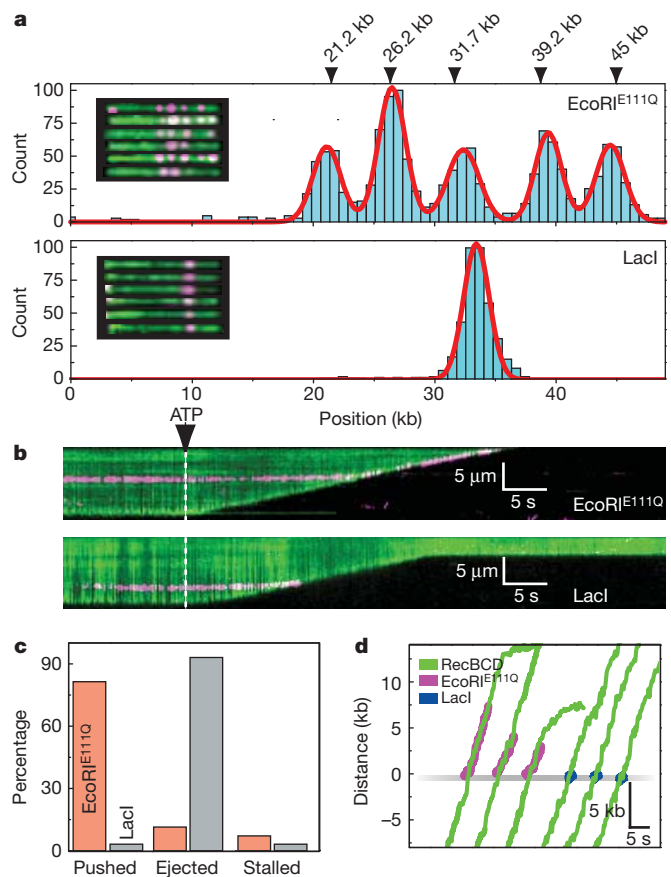


holoenzyme was heparin resistant ( $t_{1/2} \gg 6.7$  min,  $N = 58$ ), confirming open complex formation (Supplementary Fig. 3c, d). Bulk assays verified that QD–RNAP produced transcripts (Supplementary Fig. 3e), and single-molecule assays revealed a transcription velocity of  $15.7 \pm 8.6$  bp  $s^{-1}$  ( $N = 20$ ,  $25^\circ\text{C}$ ,  $250\ \mu\text{M}$  of each ribonucleoside triphosphate; Supplementary Fig. 3f).

When RecBCD collided with RNAP, the polymerase was rapidly ejected from DNA ( $t_{1/2} = 2.4 \pm 0.13$  s; Fig. 1b). Remarkably, RNAP could be pushed long distances ( $10,460 \pm 7,690$  bp,  $N = 44$ ; Fig. 1 and Supplementary Fig. 4) and RecBCD could disrupt core RNAP, holoenzymes, stalled elongation complexes and active elongation complexes (Fig. 1b and Supplementary Fig. 5). Out of 47 collisions with QD–RNAP holoenzyme, 15% (7 of 47) immediately stalled RecBCD, 8.5% (4 of 47) resulted in dissociation of RNAP with no sliding, 76.5% (36 of 47) of RNAP was pushed and 71% of pushed molecules were eventually ejected (Fig. 1c). The population of RNAP molecules that was directly ejected from the DNA increased  $\sim 5$ -fold for stalled and active elongation complexes (Supplementary Fig. 5). RecBCD also pushed and evicted RNAP labelled with 40-nm fluorescent beads or Alexa Fluor 488, arguing against nonspecific interactions between RecBCD and the quantum dots (Supplementary Fig. 6a). RecBCD did not slow or pause on colliding with RNAP (Fig. 1d and Supplementary Fig. 4a), nor was there any reduction in processivity in comparison with naked DNA ( $29,000 \pm 15,500$  bp). Similar outcomes were observed before and after Chi (not shown), indicating that RecBCD could dislodge RNAP regardless of whether RecB or RecD was the lead motor. We could unambiguously assign the orientation of RNAP at  $\lambda P_{BL}$  (Fig. 1a and Supplementary Fig. 3f), and RecBCD dislodged RNAP bound at  $\lambda P_{BL}$  during collisions in either direction (Fig. 1b and Supplementary Fig. 7a). RecBCD also pushed and ejected RNAP bound at all other locations regardless of DNA orientation (Fig. 1b). RecBCD even dislodged RNAP at lower velocities ( $446 \pm 192$  bp  $s^{-1}$ ,  $122 \pm 128$  bp  $s^{-1}$  and  $78 \pm 27$  bp  $s^{-1}$  at  $100\ \mu\text{M}$ ,  $25\ \mu\text{M}$  and  $15\ \mu\text{M}$  ATP, respectively; see Supplementary Fig. 7 and below), indicating that proteins could be dislodged even under suboptimal translocation conditions. We conclude that RecBCD disrupts RNAP regardless of orientation, transcriptional status or translocation velocity.

We next asked whether RecBCD could dislodge other proteins. EcoRI<sup>E111Q</sup> is a catalytically inactive version of EcoRI, which has high affinity ( $K_d = 2.5$  fM) for cognate sites and even binds tightly to non-specific DNA<sup>17</sup> ( $K_d = 4.8$  pM). EcoRI<sup>E111Q</sup> can halt *E. coli* RNA polymerase<sup>18,19</sup>; T7 and SP6 RNA polymerases<sup>20</sup>; SV40 large T antigen; *E. coli* UvrD, DnaB and T4 Dda helicases; SV40 replication forks<sup>21</sup>; and *E. coli* replication forks<sup>4</sup>. EcoRI withstands up to  $\sim 20$ – $40$  pN (ref. 22), and EcoRI<sup>E111Q</sup> binds cognate sites  $\sim 3000$ -fold stronger than wild-type EcoRI<sup>17</sup> ( $K_d = 6.7$  pM); thus, we infer that the catalytic mutant can resist at least as much force as the wild-type protein. *lac* repressor (LacI) is representative of a large family of bacterial transcription factors that has served as a model for transcriptional regulation and protein–DNA interactions. LacI binds tightly to specific sites<sup>23</sup> ( $K_d = 10$  fM for a 21-bp symmetric operator) but binds weakly to nonspecific DNA<sup>24</sup> ( $K_d \geq 1$  nM) and slides rapidly along nonspecific DNA rather than remaining at fixed locations<sup>25,26</sup>. LacI also blocks RNAP and replication forks both *in vitro* and *in vivo*<sup>18</sup>, highlighting that it is a potent and physiologically relevant barrier to translocase progression.

We labelled EcoRI<sup>E111Q</sup> and LacI with quantum dots (Supplementary Information), and QD–EcoRI<sup>E111Q</sup> and QD–LacI bound to the correct locations on the DNA substrates, confirming that the tagged proteins retained normal DNA-binding activity (Fig. 2a and Supplementary Fig. 8). QD–LacI was rapidly released from DNA by isopropyl- $\beta$ -D-thiogalactoside, as expected (Supplementary Fig. 9). When RecBCD collided with EcoRI<sup>E111Q</sup>, it pushed the proteins  $13,000 \pm 9,100$  bp ( $N = 70$ ) before ejecting them from the DNA (Fig. 2b, c). In contrast, LacI was immediately ejected, and was not pushed within our resolution limits (Fig. 2b–d). There was no change



**Figure 2 | Disruption of EcoRI<sup>E111Q</sup> and *lac* repressor by RecBCD.**

**a**, Histogram of EcoRI<sup>E111Q</sup> (upper panel,  $N = 1,481$ ) and LacI (lower panel,  $N = 700$ ) bound to  $\lambda$  DNA. The locations of the five EcoRI sites found in  $\lambda$  DNA are indicated, along with examples of QD–EcoRI<sup>E111Q</sup> bound to YOYO-1-stained  $\lambda$  DNA (inset, upper panel) and examples of QD–LacI bound to the DNA (inset, lower panel). **b**, Kymographs showing RecBCD colliding with EcoRI<sup>E111Q</sup> and LacI (magenta), as indicated. **d**, Distribution of event types for EcoRI<sup>E111Q</sup> and LacI. **e**, Tracking data for individual collisions.

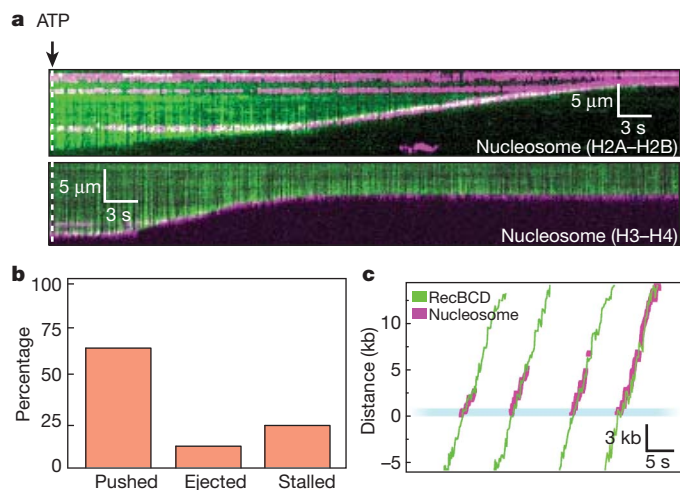
in velocity or processivity upon colliding with either protein (Fig. 2b, d and Supplementary Fig. 4). Out of 70 collisions with QD–EcoRI<sup>E111Q</sup>, 11.2% (5 of 70) stalled the translocase, 11.4% (8 of 70) resulted in immediate dissociation of EcoRI<sup>E111Q</sup> with no detectable sliding, 81.4% (57 of 70) of EcoRI<sup>E111Q</sup> was pushed along DNA and 92% of pushed molecules were eventually ejected (Fig. 2c). Out of 30 collisions with LacI, 3.3% (1 of 30) stalled the translocase, 93.3% (28 of 30) resulted in immediate dissociation of LacI with no detectable sliding and 3.3% (1 of 30) showed sliding before dissociation (Fig. 2c). A greater fraction of LacI might slide, but if so, the sliding events fall outside our resolution limits. Control experiments confirmed that RecBCD disrupted EcoRI<sup>E111Q</sup> labelled with fluorescent beads or Alexa Fluor 488 (Supplementary Fig. 6b). As with RNAP, RecBCD could strip EcoRI<sup>E111Q</sup> after Chi (not shown) and also disrupted EcoRI<sup>E111Q</sup> and LacI during low-velocity collisions (see below). These findings confirm that RecBCD readily displaces tightly bound proteins from DNA.

In eukaryotes, nucleosomes are the most frequently encountered DNA-bound obstacles. Replisomes, transcription machinery and ATP-dependent chromatin remodellers all act through mechanisms requiring force generation, and the response of nucleosomes to these forces remains a long-standing question in chromatin biology. Heterologous systems have revealed fundamental principles underlying these processes<sup>27,28</sup>: experiments with SP6 RNAP provided a theoretical framework for nucleosome repositioning<sup>27</sup>, and studies with phage T4 proteins were among the first to address the fate of nucleosomes during

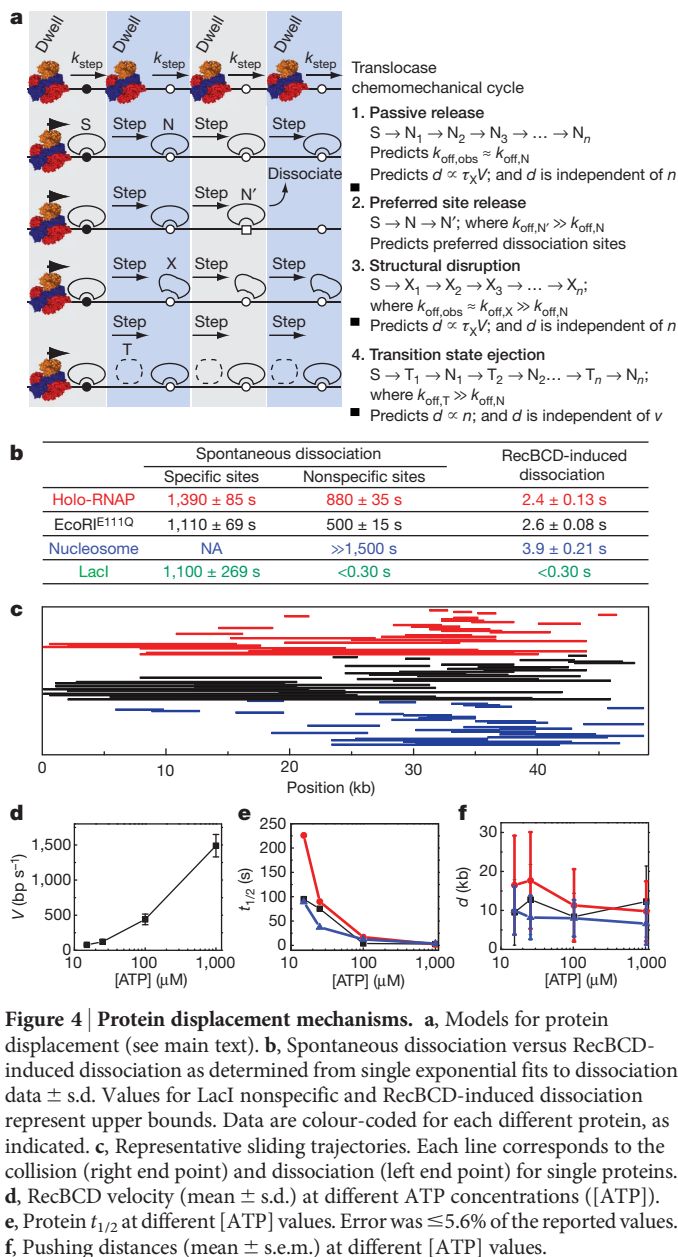
replication<sup>28</sup>. Eukaryotic translocases exert forces in the same net direction as RecBCD, and RecBCD can unwind nucleosome-bound DNA<sup>29</sup>, arguing that it can serve as a good protein-based force probe for studying the fate of nucleosomes when rammed by a translocase.

Recombinant nucleosomes were deposited on DNA curtains by salt dialysis, as described<sup>9</sup>. Remarkably, RecBCD could push nucleosomes ( $7,311 \pm 5,373$  bp,  $N = 75$ ; Fig. 3), and similar results were obtained with fluorescently labelled H2A–H2B dimer or H3–H4 tetramer (Fig. 3a). Control experiments demonstrated that RecBCD could also push nucleosomes labelled with either fluorescent beads or Alexa Fluor 488 (Supplementary Fig. 6c). Out of 357 collisions with nucleosomes, 24% (84 of 357) immediately stalled RecBCD, 11% (40 of 357) resulted in direct nucleosome ejection, 65% (233 of 357) led to sliding (Fig. 3b) and ~50% of these were eventually ejected ( $t_{1/2} = 3.93 \pm 0.21$  s; Fig. 3b and Supplementary Fig. 4c). Nucleosomes reduced the processivity of RecBCD to  $14,000 \pm 7,000$  bp, as anticipated<sup>29</sup>, and the translocase stalled in a larger fraction of these collisions (24%) than in collisions with RNAP (15%), EcoRI<sup>E111Q</sup> (7%) and LacI (3.3%). Relative to the other roadblock proteins, fewer of the pushed nucleosomes (50%) were subsequently ejected from the DNA, and there was a 10% reduction ( $t$ -test,  $P = 0.0005$ ) in velocity while pushing nucleosomes (Fig. 3c and Supplementary Fig. 4c). These results demonstrate that intact nucleosomes can be pushed along DNA as theoretically predicted<sup>30</sup>, but indicated that RecBCD had more difficulty pushing and evicting nucleosomes than it did the other protein roadblocks. The finding that RecBCD pushes and evicts nucleosomes also rules out mechanisms requiring species-specific protein–protein interactions.

Protein disruption mechanisms can be described by at least four models, which differ in the nature of the mobile intermediates and the stage of the chemomechanical cycle during which the proteins dissociate (Fig. 4a). In the first model, passive release, the proteins (S) are dislodged from a high-affinity specific site and then pushed from one sequential nonspecific site to the next. Subsequent dissociation occurs spontaneously simply because the proteins are bound to lower-affinity nonspecific DNA (N). This model assumes that the proteins have similar low affinities for all nonspecific sites sampled, and predicts that the observed rates of RecBCD-induced dissociation ( $k_{\text{off,obs}}$ ) would be similar to that of spontaneous dissociation from nonspecific DNA in the absence of RecBCD ( $k_{\text{off,obs}} \approx k_{\text{off,N}}$ ). This model also predicts that the distance ( $d$ ) over which proteins are pushed will be dictated by their affinity for nonspecific DNA and will be proportional to velocity ( $V$ ) such that faster translocation will lead to longer distances and slower



**Figure 3 | Nucleosomes can be pushed along DNA.** **a**, Kymograms showing RecBCD collisions with nucleosomes (magenta) that are labelled on either the H2A–H2B dimer or the H3–H4 tetramer, as indicated. **b**, Distribution of event types. **c**, Tracking data illustrating collisions between RecBCD and nucleosomes.



**Figure 4 | Protein displacement mechanisms.** **a**, Models for protein displacement (see main text). **b**, Spontaneous dissociation versus RecBCD-induced dissociation as determined from single exponential fits to dissociation data  $\pm$  s.d. Values for LacI nonspecific and RecBCD-induced dissociation represent upper bounds. Data are colour-coded for each different protein, as indicated. **c**, Representative sliding trajectories. Each line corresponds to the collision (right end point) and dissociation (left end point) for single proteins. **d**, RecBCD velocity (mean  $\pm$  s.d.) at different ATP concentrations ([ATP]). **e**, Protein  $t_{1/2}$  at different [ATP] values. Error was  $\leq 5.6\%$  of the reported values. **f**, Pushing distances (mean  $\pm$  s.e.m.) at different [ATP] values.

translocation will yield shorter distances. The second model, preferred site release, accounts for a situation in which proteins encounter rare sequences of exceptionally low affinity ( $N'$ ), such that they preferentially dissociate from these sites ( $k_{\text{off,N}'} \gg k_{\text{off,N}}$ ). In the third model, structural disruption, translocase collisions alter the conformation of the proteins (for example by permanently rupturing a subset of protein–DNA contacts) such that they persist as structurally perturbed complexes ( $X$ ) after displacement from the high-affinity site. In this case, the mobile intermediates have a characteristic lifetime ( $\tau_X$ ) dictated by their weakened affinity for DNA, and this lifetime should be insensitive to translocation velocity. Therefore, the distance ( $d$ ) over which proteins are pushed will be proportional to velocity ( $V$ ), and faster translocation will lead to longer distances whereas slower translocation will yield shorter distances. The most important feature of this model, which distinguishes it from all of the other models, is that the structurally disrupted proteins are more weakly bound to DNA specifically as a consequence of the collision, such that the observed rate of RecBCD-induced dissociation ( $k_{\text{off,obs}}$ ) would be greater than the rate of spontaneous dissociation from nonspecific DNA ( $k_{\text{off,obs}} \approx k_{\text{off,X}} \gg k_{\text{off,N}}$ ). The fourth model, transition state ejection, is characterized by a series of



tightly bound nonspecific complexes (N) that must pass through a weakly bound transition state (T) as they are pushed from one position to the next. This model predicts that dissociation occurs predominantly during the transition state ( $k_{\text{off},T} \gg k_{\text{off},N}$ ). The time required to pass through the transition state during one round of the chemomechanical cycle is equivalent to the time required for the translocase to take a single step ( $k_{\text{step}}$ ), which is a fixed intrinsic value independent of ATP concentration. This relationship can be rationalized by considering that the velocity of RecBCD can be controlled by modulating ATP concentration (see below), with slower velocities resulting from longer dwell times between steps (while awaiting new ATP) rather than from changes in  $k_{\text{step}}$ . Therefore, the time it takes the roadblock to pass through the transition state during a single step will be independent of ATP concentration, whereas the cumulative time spent in the transition state will increase linearly with step number ( $n$ ) irrespective of the overall observed translocation velocity. The probability of dissociation will then increase with step number, the observed lifetimes will be inversely proportional to velocity and the total distance the proteins are pushed before dissociation will be independent of velocity (that is, the roadblocks will be pushed similar distances regardless of how fast the translocase moves).

Each aforementioned model makes distinct predictions that can be experimentally evaluated. This evaluation is easier for RNAP, EcoRI<sup>E111Q</sup> and nucleosomes because these proteins are pushed long distances (LacI is considered separately below). We first measured dissociation of these proteins from specific and nonspecific sites in the absence of RecBCD (Supplementary Information), and compared these results to RecBCD-induced rates of dissociation (Fig. 4b). RNAP, EcoRI<sup>E111Q</sup> and nucleosomes all bind tightly to nonspecific DNA, and RecBCD-induced dissociation was  $\geq 200$ -fold faster than spontaneous dissociation from nonspecific sites, which is inconsistent with passive release. We next analysed pushing trajectories to determine whether there was any evidence supporting preferred site release. Comparison of these trajectories revealed that RecBCD-induced dissociation of all three roadblock proteins occurred at random locations (Fig. 4c), arguing against preferred site release. To distinguish between structural disruption and transition state eviction, we compared protein lifetimes and pushing distances at four different translocation velocities (Fig. 4d). Remarkably, a 3.3-fold decrease in RecBCD velocity ( $446 \pm 192 \text{ bp s}^{-1}$  at  $100 \mu\text{M}$  ATP) led to 1.5-, 7.0- and 3.4-fold increases in the post-collision half-lives of EcoRI<sup>E111Q</sup>, RNAP and nucleosomes (Fig. 4e), respectively, although the distribution of distances over which the proteins were pushed remained largely unaltered (Fig. 4f and Supplementary Table 1). This effect was even more obvious at  $15 \mu\text{M}$  ATP, where a 19-fold decrease in RecBCD velocity ( $78 \pm 27 \text{ bp s}^{-1}$ ) led to 36-, 93- and 24-fold increases in the post-collision half-lives of EcoRI<sup>E111Q</sup>, RNAP and nucleosomes, respectively, but pushing distances were either unaltered or increased in comparison with those corresponding to the faster velocities. These results indicated that dissociation was dictated by the number of steps the proteins were forced to take rather than the cumulative time it took to be pushed a given distance, which is most consistent with transition state ejection. Although our experiments did not reveal any evidence for a structural disruption eviction mechanism, this does not rule out the possibility that EcoRI<sup>E111Q</sup>, RNAP and nucleosomes are structurally altered when acted upon by RecBCD. However, if they are structurally perturbed, this alone does not result in their eventual dissociation from DNA.

LacI differs from the other roadblocks in that it was immediately evicted from DNA, and the RecBCD-induced dissociation rate was comparable to the rate of spontaneous dissociation from nonspecific sites (Fig. 4b), which would seem consistent with a passive-release model. However, with current resolution limits we cannot completely rule out other mechanisms, and future studies will be necessary to fully address this issue. Importantly, RNAP, EcoRI<sup>E111Q</sup> and nucleosomes all bind tightly to nonspecific DNA, whereas LacI binds much more weakly to nonspecific sequences (Fig. 4b), suggesting that LacI is released more

rapidly from DNA after the collisions due to its weaker affinity for nonspecific sites. This result demonstrates that the roadblock proteins and the nature of their interactions with nonspecific DNA are critical contributing factors to the outcome of the collisions.

This leaves the question of how much force RecBCD exerts, and how much is sufficient to disrupt obstacles. Although our experiments do not yield a direct read-out of force, we can safely conclude that the force exerted by RecBCD is sufficient to displace RNAP, EcoRI<sup>E111Q</sup> LacI and nucleosomes from DNA. Our work has revealed unprecedented details of protein collisions on DNA and provides new insights into how translocases can disrupt nucleoprotein complexes. Given the flexibility of our experimental platform, we anticipate that these studies can be extended to other translocases and roadblock proteins, and it will be important to determine whether the mechanistic concepts developed here apply to different types of collision between proteins on DNA.

## METHODS SUMMARY

We conducted total-internal-reflection fluorescence microscopy experiments on a home-built microscope using nanofabricated DNA curtains, as previously described<sup>9</sup>. For all initial experiments, and for all kymographs shown in the manuscript, we used YOYO-1 to stain the DNA. YOYO-1 does not affect the translocation rate or processivity of RecBCD<sup>6</sup>, and it did not affect the binding distributions of RNAP, EcoRI<sup>E111Q</sup> or nucleosomes (not shown). In the presence of YOYO-1, the roadblock proteins showed the same general response to collisions with RecBCD, with identical distributions of ejection, stalling and pushing (and pushing velocities) seen with and without YOYO-1. However, the stain reduced the distance obstacles were pushed by 20–30%. Therefore, all sliding distances and half-lives reported here correspond to values measured in the absence of YOYO-1. Sliding distances are reported only for roadblock proteins that did not encounter any other quantum-dot-tagged proteins as they were pushed along the DNA. This ensures that each analysed collision/dissociation event involved only a single quantum-dot-tagged protein. Many reactions were observed in which multiple quantum-dot-tagged roadblock proteins were pushed into one another, but in these cases we could not determine the order in which each such protein was displaced from the DNA, and therefore could not measure sliding distances. To categorize the distributions of event type, we defined 'sliding' as the movement of any quantum-dot-tagged roadblock by more than  $0.53 \mu\text{m}$  ( $\sim 1,950 \text{ bp}$ ); anything less than this was scored as a direct dissociation event.

Received 8 August; accepted 6 October 2010.

Published online 24 November 2010.

1. Jankowsky, E., Gross, C., Shuman, S. & Pyle, A. Active disruption of an RNA-protein interaction by a DExH/D RNA helicase. *Science* **291**, 121–125 (2001).
2. Marquis, K. A. *et al.* SpoIIIE strips proteins off the DNA during chromosome translocation. *Genes Dev.* **22**, 1786–1795 (2008).
3. Krejci, L. *et al.* DNA helicase Srs2 disrupts the Rad51 presynaptic filament. *Nature* **423**, 305–309 (2003).
4. Guy, C. P. *et al.* Rep provides a second motor at the replisome to promote duplication of protein-bound DNA. *Mol. Cell* **36**, 654–666 (2009).
5. Singleton, M. R., Dillingham, M., Gaudier, M., Kowalczykowski, S. & Wigley, D. Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks. *Nature* **432**, 187–193 (2004).
6. Bianco, P. R. *et al.* Processive translocation and DNA unwinding by individual RecBCD enzyme molecules. *Nature* **409**, 374–378 (2001).
7. Spies, M., Amitani, I., Baskin, R. & Kowalczykowski, S. RecBCD enzyme switches lead motor subunits in response to chi recognition. *Cell* **131**, 694–705 (2007).
8. Taylor, A. F. & Smith, G. R. RecBCD enzyme is a DNA helicase with fast and slow motors of opposite polarity. *Nature* **423**, 889–893 (2003).
9. Visnapuu, M.-L. & Greene, E. Single-molecule imaging of DNA curtains reveals intrinsic energy landscapes for nucleosome deposition. *Nature Struct. Mol. Biol.* **16**, 1056–1062 (2009).
10. Ishihama, A. Functional modulation of *Escherichia coli* RNA polymerase. *Annu. Rev. Microbiol.* **54**, 499–518 (2000).
11. Herbert, K. M., Greenleaf, W. J. & Block, S. M. Single-molecule studies of RNA polymerase: motoring along. *Annu. Rev. Biochem.* **77**, 149–176 (2008).
12. Liu, B., Wong, M. & Alberts, B. A transcribing RNA polymerase molecule survives DNA replication without aborting its growing RNA chain. *Proc. Natl Acad. Sci. USA* **91**, 10660–10664 (1994).
13. Liu, B., Wong, M., Tinker, R., Geiduschek, E. & Alberts, B. The DNA replication fork can pass RNA polymerase without displacing the nascent transcript. *Nature* **366**, 33–39 (1993).
14. Liu, B. & Alberts, B. Head-on collision between a DNA replication apparatus and RNA polymerase transcription complex. *Science* **267**, 1131–1137 (1995).



15. Pomerantz, R. T. & O'Donnell, M. The replisome uses mRNA as a primer after colliding with RNA polymerase. *Nature* **456**, 762–766 (2008).
16. Pomerantz, R. T. & O'Donnell, M. Direct restart of a replication fork stalled by a head-on RNA polymerase. *Science* **327**, 590–592 (2010).
17. Wright, D. J., King, K. & Modrich, P. The negative charge of Glu-111 is required to activate the cleavage center of EcoRI endonuclease. *J. Biol. Chem.* **264**, 11816–11821 (1989).
18. Epshtein, V. & Toulm, E. F. Rahmouni, A. Borukhov, S. & Nudler, E. Transcription through the roadblocks: the role of RNA polymerase cooperation. *EMBO J.* **22**, 4719–4727 (2003).
19. Nudler, E., Kashlev, M., Nikiforov, V. & Goldfarb, A. Coupling between transcription termination and RNA polymerase inchworming. *Cell* **81**, 351–357 (1995).
20. Pavco, P. A. & Steege, D. A. Characterization of elongating T7 and SP6 RNA polymerases and their response to a roadblock generated by a site-specific DNA binding protein. *Nucleic Acids Res.* **19**, 4639–4646 (1991).
21. Byrd, A. K. & Raney, K. D. Displacement of a DNA binding protein by Dda helicase. *Nucleic Acids Res.* **34**, 3020–3029 (2006).
22. Noom, M. C., van den Broek, B., van Mameren, J. & Wuite, G. J. L. Visualizing single DNA-bound proteins using DNA as a scanning probe. *Nature Methods* **4**, 1031–1036 (2007).
23. Sadler, J. R., Sasmor, H. & Betz, J. L. A perfectly symmetric lac operator binds the lac repressor very tightly. *Proc. Natl Acad. Sci. USA* **80**, 6785–6789 (1983).
24. Lin, S.-Y. & Riggs, A. D. Lac repressor binding to DNA not containing the lac operator and to synthetic poly dAT. *Nature* **228**, 1184–1186 (1970).
25. Elf, J., Li, G.-W. & Xie, X. Probing transcription factor dynamics at the single-molecule level in a living cell. *Science* **316**, 1191–1194 (2007).
26. Wang, Y. M., Austin, R. H. & Cox, E. C. Single molecule measurements of repressor protein 1D diffusion on DNA. *Phys. Rev. Lett.* **97**, 048302 (2006).
27. Studitsky, V. M., Clark, D. J. & Felsenfeld, G. Overcoming a nucleosomal barrier to transcription. *Cell* **83**, 19–27 (1995).
28. Bonne-Andrea, C., Wong, M. & Alberts, B. *In vitro* replication through nucleosomes without histone displacement. *Nature* **343**, 719–726 (1990).
29. Eggleston, A. K., O'Neill, T. E., Bradbury, E. M. & Kowalczykowski, S. C. Unwinding of nucleosomal DNA by a DNA helicase. *J. Biol. Chem.* **270**, 2024–2031 (1995).
30. Mollazadeh-Beidokhti, L., Deseigne, J., Lacoste, D., Mohammad-Rafiee, F. & Schiessel, H. Stochastic model for nucleosome sliding under an external force. *Phys. Rev. E* **79**, 031922 (2009).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank M. Gottesman, R. Gonzalez and members of the Greene laboratory for discussion and assistance throughout this work. We thank P. Modrich for providing an expression construct encoding EcoRI<sup>E111Q</sup>, R. Landick and K. Adelman for providing RNAP constructs, and J. Gelles for providing plasmids encoding RecBCD. I.J.F. was supported by an NIH Fellowship (F32GM80864). Funding was provided by the National Institutes of Health (GM074739 and GM082848 to E.C.G.). This work was partially supported by the Initiatives in Science and Engineering program through Columbia University, the Nanoscale Science and Engineering Initiative of the National Science Foundation under NSF Award Number CHE-0641523, and by the New York State Office of Science, Technology, and Academic Research. E.C.G. is an Early Career Scientist with the Howard Hughes Medical Institute. We apologize to colleagues whose work we were unable to cite owing to length restrictions.

**Author Contributions** I.J.F. did all cloning and ensemble-level biochemical characterization, and conducted and analysed RecBCD collision experiments with RNAP, EcoRI<sup>E111Q</sup> and LacI. M.-L.V. conducted and analysed RecBCD collision experiments with nucleosomes. I.J.F., M.-L.V. and E.C.G. discussed the data and co-wrote the paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to E.C.G. ([ecg2108@columbia.edu](mailto:ecg2108@columbia.edu)).

# The mechanism of sodium and substrate release from the binding pocket of vSGLT

Akira Watanabe<sup>1\*</sup>, Seungho Choe<sup>2\*</sup>, Vincent Chaptal<sup>1</sup>, John M. Rosenberg<sup>2,3</sup>, Ernest M. Wright<sup>1</sup>, Michael Grabe<sup>2,3</sup> & Jeff Abramson<sup>1</sup>

**Membrane co-transport proteins that use a five-helix inverted repeat motif have recently emerged as one of the largest structural classes of secondary active transporters<sup>1,2</sup>. However, despite many structural advances there is no clear evidence of how ion and substrate transport are coupled. Here we report a comprehensive study of the sodium/galactose transporter from *Vibrio parahaemolyticus* (vSGLT), consisting of molecular dynamics simulations, biochemical characterization and a new crystal structure of the inward-open conformation at a resolution of 2.7 Å. Our data show that sodium exit causes a reorientation of transmembrane helix 1 that opens an inner gate required for substrate exit, and also triggers minor rigid-body movements in two sets of transmembrane helical bundles. This cascade of events, initiated by sodium release, ensures proper timing of ion and substrate release. Once set in motion, these molecular changes weaken substrate binding to the transporter and allow galactose readily to enter the intracellular space. Additionally, we identify an allosteric pathway between the sodium-binding sites, the unwound portion of transmembrane helix 1 and the substrate-binding site that is essential in the coupling of co-transport.**

Secondary active transporters harness the energy stored in electrochemical gradients to drive the accumulation of specific solutes across cell membranes. This task is accomplished by the alternating-access mechanism, in which the substrate-binding site is first exposed to one side of the membrane and, on ion and substrate binding, a conformational change exposes the transported solute to the opposite face, where it is released<sup>3</sup>. Sodium/glucose co-transporters are prototypes of secondary active transporters that drive the accumulation of sugars and other molecules into cells. These transporters have critical roles in human physiology, where mutations in their genes are responsible for severe congenital diseases<sup>4</sup> and are the molecular targets for drugs to treat diabetes and obesity<sup>5</sup>.

There has been a recent surge of work on crystal structures<sup>6–11</sup> displaying the five-helix inverted repeat motif. These are referred to as the LeuT superfamily and include genetically diverse proteins that transport a wide range of substrates and differ in the number and type of driving ligand<sup>1,2,12</sup>. A general model for alternating access is being pieced together through comparisons of these diverse structures<sup>1,2,13</sup>. Despite sharing a common set of ten core transmembrane segments, the lack of sequence similarity and the chemical diversity of the transported substrates prevents the complete understanding of the mechanistic basis of transport. This hurdle is being surmounted as multiple structures of the same protein—at different stages in the transport cycle—are solved, providing a comprehensive understanding of substrate binding<sup>7,14–16</sup> and the transition from outward- to inward-facing conformations<sup>17</sup>. However, an atomic-level understanding of sodium-coupled substrate co-transport, necessary to explain the dynamics of alternating access, is still absent.

To investigate the mechanism of sodium–sugar coupling, we carried out molecular dynamics simulations on the galactose-bound inward-occluded conformation of vSGLT<sup>6</sup> embedded in a lipid bilayer<sup>18</sup>. All sodium co-transporters of the LeuT superfamily share a common

sodium-binding site termed the Na2 site. During the transition to the inward-facing conformation, transmembrane helix (TM) 8, which forms part of the sodium-binding site, is displaced by ~4 Å, generating a less favourable Na2 site that facilitates Na<sup>+</sup> release<sup>1,17</sup> (Fig. 1b). Na<sup>+</sup> modelled at this site is loosely coordinated by the carbonyl oxygens of Ile 65 (3.3 Å), Ala 361 (3.2 Å) and the side-chain hydroxyl of Ser 365 (3.1 Å). The carbonyl oxygen of Ala 62 (3.6 Å) and the side-chain hydroxyl of Ser 364 (3.6 Å) are also in close proximity (Fig. 1b). Previous molecular dynamics simulations performed on vSGLT<sup>19</sup> and Mhp1<sup>17</sup> indicated that Na<sup>+</sup> quickly leaves the Na2 site. Our simulations indicate that Na<sup>+</sup> exits the Na2 site after 9 ns (Fig. 2a) and interacts with the hydrophilic pore-lining residue Asp 189 on TM5 during exit. The importance of Asp 189 was highlighted in a previous simulation<sup>19</sup> and in biochemical studies on hSGLT<sup>20</sup>. All three molecular dynamics simulations indicate that Na<sup>+</sup> exits the transporter before substrate exit; however, additional conformational changes are required to release the occluded galactose.

In the inward-occluded structure, galactose is located halfway across the membrane (Figs 1a and 2b), where it is coordinated by extensive side-chain interactions from TM1, TM2, TM6, TM7 and TM10. Subsets of these residues form two hydrophobic gates blocking galactose exit to the intracellular and extracellular spaces. Our molecular dynamics simulations show that as Na<sup>+</sup> exits the Na2 site, galactose undergoes significant fluctuations within the binding pocket. At 52 ns, Tyr 263 adopts a new and stable rotamer conformation that expands the exit pathway (between 52 and 110 ns), permitting the sugar to leave the binding site (Figs 2 and 3). After sugar release (~110 ns), Tyr 263 returns to the original conformation.

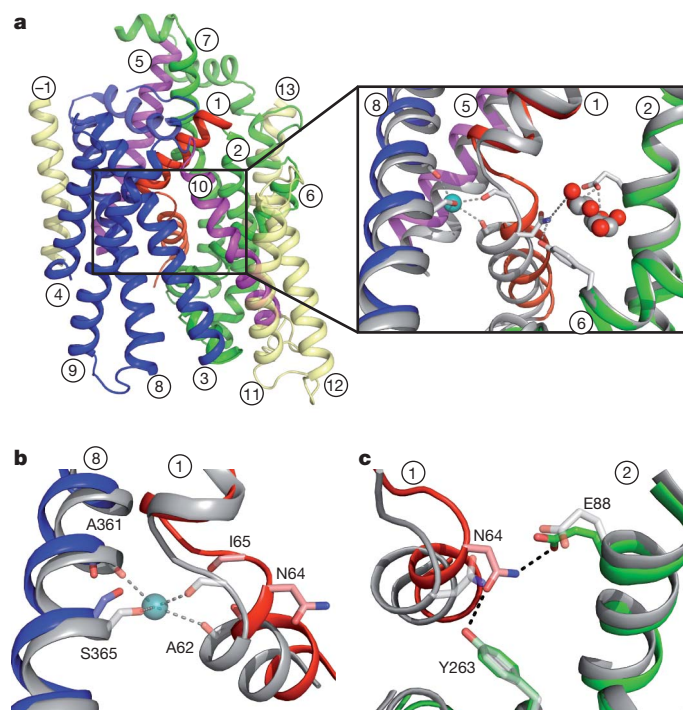
To test the hypothesis that Na<sup>+</sup> release stimulates an alternative conformation of Tyr 263, we conducted a 200-ns molecular dynamics simulation in which the sodium was lightly restrained in the Na2 site. Under these conditions, Tyr 263 never adopts the alternative conformation, and thereby prevents galactose exit (Supplementary Fig. 1). This observation suggests that sodium release drives conformational changes that disrupt the galactose-binding site, and further suggests that interactions between the Na2 site and Tyr 263 are central to the transport mechanism.

The spontaneous release of galactose in the absence of applied forces makes possible the accurate determination of the binding free energy profile through the use of umbrella sampling along the exit pathway coupled with weighted histogram analysis<sup>21</sup> (Fig. 3, inset). After Na<sup>+</sup> release, galactose is weakly bound to vSGLT with a minimal energy barrier of ~2 kcal mol<sup>-1</sup>, resulting from the interaction of the sugar with residues Asn 64, Ser 66, Glu 68 and Gln 69 on TM1. Asn 64 is of particular interest because it is located in the unwound segment of TM1 and has hydrogen bonds with the inner gate residue Tyr 263 and the O2 hydroxyl of galactose linking the Na2 site with the galactose site. Thus, the interactions of Asn 64 with Tyr 263 and galactose may be critical to the transport mechanism<sup>6</sup>.

To test the importance of these interactions, we performed molecular dynamics simulations and sodium-dependent transport assays on

<sup>1</sup>Department of Physiology, University of California, Los Angeles, Los Angeles, California 90095-1759, USA. <sup>2</sup>Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA. <sup>3</sup>Department of Computational and Systems Biology, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA.

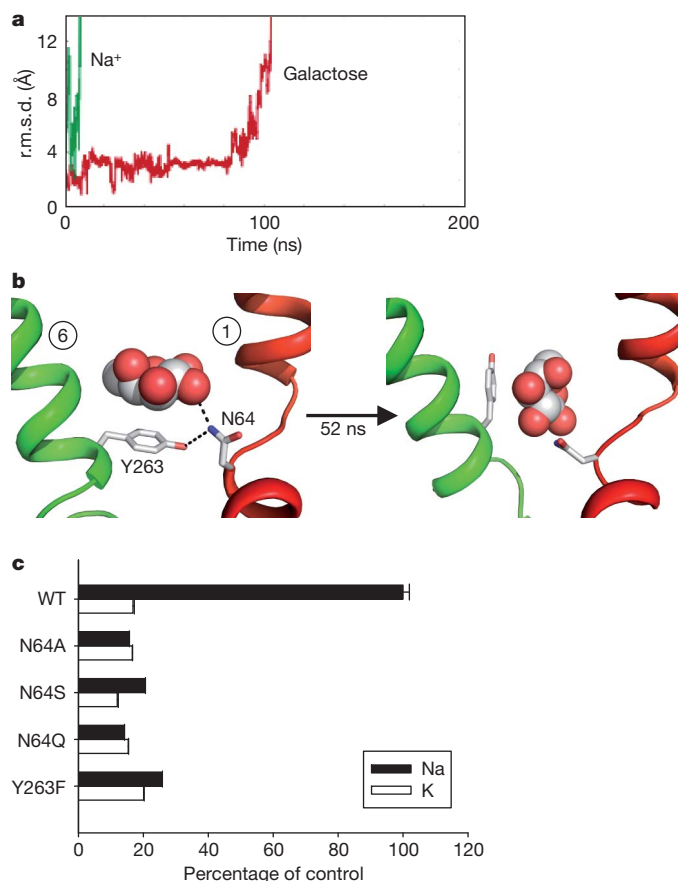
\*These authors contributed equally to this work.



**Figure 1 | Structures and overlay of the inward-open and inward-occluded conformations.** **a**, The core domain of the inward-open conformation (TM1–TM10) is coloured by specific helix bundles involved in the transition from the inward-occluded to the inward-open conformation. The ‘hash motif’ formed from TM3, TM4, TM8 and TM9 is blue; the ‘sugar bundle’ formed from TM2, TM6 and TM7 is green; TM1 is red; and TM5 and TM10 are magenta. The periphery helices (TM–1, TM11, TM12 and TM13) are yellow. Atoms are displayed in ball-and-stick form with oxygen coloured red and nitrogen coloured blue. Inset, an overlay of the inward-open (colour) and inward-occluded (grey) conformations illustrating the coordination at the Na2 and galactose-binding sites. **b**, **c**, Overlay of the inward-open and inward-occluded conformations with the same colouring as in **a**. Conformational changes in the inward-open structure reveals a  $\sim 13^\circ$  kink in the unwound segment of TM1 that prevents sodium coordination at the Na2 site (**b**). In the absence of galactose, the galactose-binding residue Asn 64 hydrogen-bonds to Glu 88 and Tyr 263, maintaining an open pathway from the intracellular space to the substrate-binding site (**c**).

transporters with mutations at positions 64 and 263. Simulations of the Asn 64 Ala mutant show a momentary sodium departure from the Na2 site at 5 ns, but the ion rapidly returns and remains for the remainder of the simulation. The failure of  $\text{Na}^+$  to unbind prevents conformational changes in the unwound segment of TM1, and Tyr 263 remains in the blocked orientation (Supplementary Fig. 2a). In agreement with the simulation, sodium-dependent transport assays on the Asn 64 Ala mutant show no activity (Fig. 2c).

To explore the role of Asn 64 further, we tested Asn 64 Gln and Asn 64 Ser, which, in principal, are both capable of maintaining the native hydrogen bonds to Tyr 263 and galactose. Models of Asn 64 Gln prevented simulation as the result of substantial steric clashes, which correlated well with a lack of transport (Fig. 2c). The model of the Asn 64 Ser mutation could form a hydrogen bond with Tyr 263 (3.3 Å) but not with the O2 hydroxyl of galactose (4.3 Å). In the simulation, Asn 64 Ser releases  $\text{Na}^+$  from the Na2 site at 25 ns, and Tyr 263 transiently adopts the alternative rotamer conformation before returning to its original position, preventing galactose exit (Supplementary Fig. 2c). Similarly, simulation of the Tyr 263 Phe mutant shows that  $\text{Na}^+$  unbinds at 10 ns, but unlike tyrosine, phenylalanine never adopts a conformation compatible with galactose exit (Supplementary Fig. 2b). Longer simulations may reveal galactose release in these mutants, because they both show modest transport activity ( $\sim 10\%$  of wild type); however, the transport assays and simulation data demonstrate that



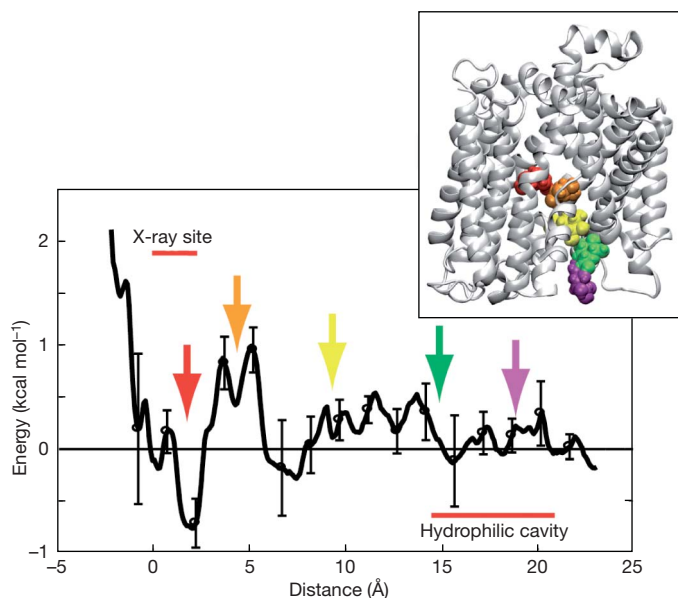
**Figure 2 | Mechanism of galactose release.** **a**, Sodium and galactose exit vSGLT. The root mean squared deviation (r.m.s.d.) of  $\text{Na}^+$  (green) rapidly increases at 9 ns, indicating exit from the Na2 site. This is followed by the release of galactose (red) at 110 ns. **b**, Tyr 263 adopts two rotamers. On the left, Tyr 263 is shown in the conformation observed in the inward-occluded structure<sup>6</sup>, in which it blocks substrate exit through a hydrogen bond with Asn 64 on TM1. At 52 ns (shown on the right), Tyr 263 adopts a rotamer conformation that expands the exit pathway. **c**, D-galactose uptake by wild-type and vSGLT mutants in proteoliposomes. Results are expressed as percentage uptake in either 100 mM NaCl or KCl, and show that the mutants Asn 64 Ala, Asn 64 Ser, Asn 64 Gln and Tyr 263 Phe severely impair sodium-dependent transport. Error bars, s.e.m. WT, wild type.

robust transport requires precise orientation of Asn 64 to stabilize galactose and the gating residue Tyr 263.

Although the molecular dynamics simulations and biochemical studies demonstrate a physical link between the Na2 site and the substrate, global details regarding the inward-open conformation (devoid of both ligands) remain elusive. To address this issue, we determined the structure of vSGLT in the inward-open conformation. Crystals, in the absence of ligands, for both the wild-type protein and the inactive Lys 294 Ala mutant<sup>6</sup> were obtained. Both crystals had the same overall configuration, but the mutant crystals diffracted to a higher resolution (2.7 Å; see Methods).

As in the original structure<sup>6</sup>, the inward-open conformation is composed of 14 transmembrane helices, ten of which comprise the core domain. TM1–TM5 and TM6–TM10 are related by an approximate two-fold symmetry axis through the centre of the membrane plane. The inward-occluded and the inward-open structures have a similar overall fold with a r.m.s.d. of 1.2 Å. However, there are distinct structural differences between the two conformations, presumably owing to changes resulting from the release of ligands (Fig. 4). With the exception of TM1, superimpositions of individual helices reveal that the occluded-to-open transition occurs by rigid-body movements of sub-domains (Supplementary Figs 3 and 4). Consistent with the recent



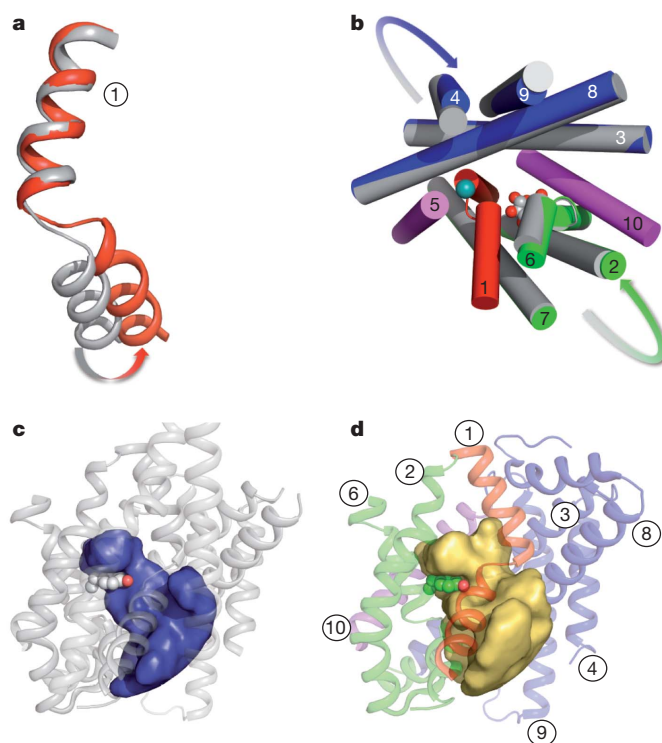


**Figure 3 | The potential of mean force for galactose unbinding.** Energy of galactose binding to vSGLT in the absence of  $\text{Na}^+$ . Umbrella sampling along the natural, equilibrium pathway shown (inset) was used to determine the binding free energy. The distance along the pathway from the binding site in the X-ray structure is shown along the  $x$  axis. The coloured arrows correspond to the galactose positions shown in the inset. The largest barrier is  $\sim 2 \text{ kcal mol}^{-1}$ , at 5 Å, which corresponds to galactose interaction with residues in the kink region of TM1. Error bars were determined by splitting the production data into four equal sets, computing the energy profile for each set, and then applying a global shift to each curve before calculating the standard deviation at the 16 positions marked with points.

assignment for Mhp1<sup>17</sup>, the hash motifs, formed from TM3 and TM4 and their inverted repeat equivalents, TM8 and TM9, align with a r.m.s.d. of 0.9 Å. TM2, TM6 and TM7 form a domain termed the sugar bundle for the extensive side-chain interactions with galactose, and these regions superimpose with a r.m.s.d. of 0.5 Å. This new inward-open structure of vSGLT is more similar to the recent inward-facing conformation structure of Mhp1 than is the previous structure of vSGLT. Details of this structural analysis are in Methods.

The transition from the inward-occluded to the inward-open structure is presumably triggered by sodium release from the Na2 site and the alteration in the hydrogen-bonding network surrounding the unwound segment of TM1. In particular, the intracellular half of TM1 flexes  $\sim 13^\circ$ , modifying the coordination of Asn 64 (Figs 1b and 4a). In the absence of both galactose and  $\text{Na}^+$ , Asn 64 coordinates Tyr 263 and Glu 88. Glu 88 was previously hydrogen-bonded to the O2 and O3 hydroxyls of galactose. This new conformation of TM1 is further stabilized by hydrogen bonds between the Na2-site residue Ser 365 and Glu 68 on the unwound segment of TM1 (Supplementary Fig. 5). When viewed from the intracellular side, each domain moves  $\sim 3^\circ$  in opposite directions, thereby increasing the volume of the accessibility cavity by  $\sim 1,400 \text{ Å}^3$  (Fig. 4). This  $6^\circ$  relative rotation probably disrupts protein–substrate coordination and permits water to enter the site. This interpretation is supported by our simulations, in which an increase in the number of water molecules in the substrate-binding site is observed after sodium release (Supplementary Fig. 6). Water effectively competes with the protein for hydrogen bonds, loosening galactose in the pocket and ultimately assisting in its release (Supplementary Fig. 6 and Supplementary Movie 1).

We propose the following mechanism for sodium and galactose exit from vSGLT. The transition from the outward- to the inward-occluded conformation weakens the Na2 sodium-binding site, causing it to become metastable and release the ion on a short timescale. Upon exit, the carbonyl oxygens of the ion-coordinating residues Ile 65 and



**Figure 4 | Conformational changes in the transition from the inward-occluded to the inward-open structure.** **a**, TM1 superimposed between the inward-open (red) and inward-occluded (grey) structures, showing a  $\sim 13^\circ$  kink in TM1. **b**, Overlay of the inward-open (coloured as in Fig. 1) and inward-occluded (grey) conformations. Rigid-body rotations of the hash motif and sugar bundle by  $3^\circ$  in opposite directions expose the substrate-binding site to the intracellular environment. **c**, Accessibility cavity of the inward-occluded conformation is coloured blue. **d**, Accessibility cavity of the inward-open conformation is coloured gold. The conformational changes from TM1, hash motif and sugar bundle cause an increase of  $\sim 1,400 \text{ Å}^3$  in the accessible volume of the inward-open conformation, aiding galactose release.

Ala 62 undergo a conformational change in the unwound segment of TM1, producing a kink of  $\sim 13^\circ$  (Figs 1 and 4a). Our simulation shows that movement of TM1 disrupts the hydrogen bond between Asn 64 and Tyr 263, allowing the side chain of Tyr 263 to adopt a new conformation that opens a pathway to the intracellular space (Fig. 2b). Additional rigid-body movements widen the intracellular cavity, allowing water penetration and further disrupting the substrate-binding site to enhance exit and prevent rebinding (Fig. 4).

It is likely that the reaction scheme described here for vSGLT is broadly used by all sodium-dependent members of the LeuT superfamily, because the Na2 site, the hydrophobic gates and the unwound segments on TM1 and TM6 are all conserved<sup>11,12,22,23</sup>. For proteins with a single sodium-binding site, the Na2 site directly interacts with the substrate through polar residues—Asn 64 in vSGLT and Gln 42 in Mhp1—located on the unwound segment of TM1. For proteins that harbour two sodium-binding sites, such as LeuT and, putatively, BetP, the additional site (the Na1 site) is positioned on the opposite side of the unwound helix from the Na2 site. Interactions between the Na1 and Na2 sites are mediated by the unwound segment of TM1, and the sodium at the Na1 site is directly coordinated to the substrate<sup>8,12</sup>. Regardless of whether the protein has one or two sodium-binding sites, it is the conserved Na2 site, positioned most distal from the core of the protein, that regulates sodium and substrate release. This primary structural feature coupling sodium and substrate co-transport has fundamental implications for our understanding of membrane protein biology and for developing strategies to manipulate the alternating-access mechanism therapeutically.

## METHODS SUMMARY

**Molecular dynamics simulations.** The vSGLT monomer (Protein Data Bank ID, 3DH4) was embedded and solvated in a 1-palmitoyl-2-oleoyl phosphatidylcholine membrane bilayer using the OPM<sup>24</sup> and CHARMM-GUI<sup>18</sup> software packages. Simulations were carried out using NAMD<sup>25</sup> with the CHARMM27 parameter set in a 150 mM NaCl bath. See Methods for more details.

**Protein expression and purification.** Plasmids carrying wild-type or mutant transporters were transformed and overexpressed in the TOP10 *Escherichia coli* cell line. Cell membranes were isolated, solubilized (2% w/v decyl- $\beta$ -D-maltopyranoside) and tandem-purified using a Ni-NTA Superflow column (affinity chromatography) and a Superdex 200 column (size-exclusion chromatography). See Methods for more details.

**Transport assays.** We generated proteoliposomes by reconstituting purified vSGLT protein with sonicated lipid at a protein/lipid ratio of 1:200. We measured transport activity by monitoring the uptake of D-galactose, with <sup>14</sup>C-D-galactose tracer, into proteoliposomes in the presence or absence of a 100 mM Na<sup>+</sup> gradient (K<sup>+</sup> replacing Na<sup>+</sup>). See Methods for more details.

**Crystallization and data collection.** We concentrated purified wild-type and Lys 294 Ala protein to ~13 mg ml<sup>-1</sup> and grew crystals by the hanging-drop vapour diffusion method using the Mosquito nanolitre-dispensing robot. Data collected at the Advanced Light Source, Berkeley (beamline 5.0.2), were integrated and scaled, and phases were calculated by molecular replacement. The model was built and refined to an  $R_{\text{work}}/R_{\text{free}}$  value of 25.1/27.4. See Methods for more details.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 11 May; accepted 12 October 2010.

Published online 5 December 2010.

1. Abramson, J. & Wright, E. M. Structure and function of Na<sup>+</sup>-symporters with inverted repeats. *Curr. Opin. Struct. Biol.* **19**, 425–432 (2009).
2. Krishnamurthy, H., Piscitelli, C. L. & Gouaux, E. Unlocking the molecular secrets of sodium-coupled transporters. *Nature* **459**, 347–355 (2009).
3. Jardetzky, O. Simple allosteric model for membrane pumps. *Nature* **211**, 969–970 (1966).
4. Wright, E. M., Hirayama, B. A. & Loo, D. F. Active sugar transport in health and disease. *J. Intern. Med.* **261**, 32–43 (2007).
5. Isaji, M. Sodium-glucose cotransporter inhibitors for diabetes. *Curr. Opin. Investig. Drugs* **8**, 285–292 (2007).
6. Faham, S. *et al.* The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na<sup>+</sup>/sugar symport. *Science* **321**, 810–814 (2008).
7. Weyand, S. *et al.* Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. *Science* **322**, 709–713 (2008).
8. Ressel, S., Terwisscha van Scheltinga, A. C., Vonrhein, C., Ott, V. & Ziegler, C. Molecular basis of transport and regulation in the Na<sup>+</sup>/betaine symporter BetP. *Nature* **458**, 47–52 (2009).
9. Fang, Y. *et al.* Structure of a prokaryotic virtual proton pump at 3.2 Å resolution. *Nature* **460**, 1040–1043 (2009).
10. Gao, X. *et al.* Structure and mechanism of an amino acid antiporter. *Science* **324**, 1565–1568 (2009).

11. Shaffer, P. L., Goehring, A., Shankaranarayanan, A. & Gouaux, E. Structure and mechanism of a Na<sup>+</sup>-independent amino acid transporter. *Science* **325**, 1010–1014 (2009).
12. Yamashita, A., Singh, S. K., Kawate, T., Jin, Y. & Gouaux, E. Crystal structure of a bacterial homologue of Na<sup>+</sup>/Cl<sup>-</sup>-dependent neurotransmitter transporters. *Nature* **437**, 215–223 (2005).
13. Forrest, L. R. & Rudnick, G. The rocking bundle: a mechanism for ion-coupled solute flux by symmetrical transporters. *Physiology (Bethesda)* **24**, 377–386 (2009).
14. Zhou, Z. *et al.* LeuT-desipramine structure reveals how antidepressants block neurotransmitter reuptake. *Science* **317**, 1390–1393 (2007).
15. Gao, X. *et al.* Mechanism of substrate recognition and transport by an amino acid antiporter. *Nature* **463**, 828–832 (2010).
16. Singh, S. K., Piscitelli, C. L., Yamashita, A. & Gouaux, E. A competitive inhibitor traps LeuT in an open-to-out conformation. *Science* **322**, 1655–1661 (2008).
17. Shimamura, T. *et al.* Molecular basis of alternating access membrane transport by the sodium-hydantoin transporter Mhp1. *Science* **328**, 470–473 (2010).
18. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865 (2008).
19. Li, J. & Tajkhorshid, E. Ion-releasing state of a secondary membrane transporter. *Biophys. J.* **97**, L29–L31 (2009).
20. Quick, M., Loo, D. D. & Wright, E. M. Neutralization of a conserved amino acid residue in the human Na<sup>+</sup>/glucose transporter (hSGLT1) generates a glucose-gated H<sup>+</sup> channel. *J. Biol. Chem.* **276**, 1728–1734 (2001).
21. Kumar, S., Bouzida, D., Swendsen, R. H., Kollman, P. A. & Rosenberg, J. M. The weighted histogram analysis method for free-energy calculations on biomolecules. 1. The method. *J. Comput. Chem.* **13**, 1011–1021 (1992).
22. Yernool, D., Boudker, O., Jin, Y. & Gouaux, E. Structure of a glutamate transporter homologue from *Pyrococcus horikoshii*. *Nature* **431**, 811–818 (2004).
23. Hunte, C. *et al.* Structure of a Na<sup>+</sup>/H<sup>+</sup> antiporter and insights into mechanism of action and regulation by pH. *Nature* **435**, 1197–1202 (2005).
24. Lomize, A. L., Pogozheva, I. D., Lomize, M. A. & Mosberg, H. I. Positioning of proteins in membranes: a computational approach. *Protein Sci.* **15**, 1318–1333 (2006).
25. Kale, L. *et al.* NAMD2: Greater scalability for parallel molecular dynamics. *J. Comput. Phys.* **151**, 283–312 (1999).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank T. Vondriska and K. Philipson as well as members of the Abramson, Wright and Grabe labs for useful discussions and for reading the manuscript. We would also like to thank S. Faham for contributions at the early stages of this work, S. Iwata for advance release of the Mhp1 coordinates (Protein Data Bank ID, 2X79), and R. Roskies for assistance with the computations. Simulations were carried out through a TeraGrid grant at the Pittsburgh Supercomputing Center and the Texas Advanced Computing Center. This work was supported by NIH grants GM078844 (J.A.), RGY0069 (J.A.) and DK19567 (E.M.W.), and a grant from the Human Frontier Science Program (J.A.). M.G. is an Alfred P. Sloan Research Fellow.

**Author Contributions** Experiments were carried out and diffraction data collected by A.W., V.C. and J.A. Simulations were carried out by S.C. Data were analysed and the manuscript was prepared by all authors.

**Author Information** Coordinates and structure factors of the inward-open vSGLT structure have been deposited in the Protein Data Bank under accession number 2XQ2. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to J.A. ([jabramson@mednet.ucla.edu](mailto:jabramson@mednet.ucla.edu)) or M.G. ([mdgrabe@pitt.edu](mailto:mdgrabe@pitt.edu)).

## METHODS

**Molecular dynamics simulations.** Initially, TM-1 was removed and six missing residues in the TM4-TM5 loop were added with the loop modelling routine in Modeller<sup>26</sup>. Residues 53–547 were then embedded in a membrane and solvated in a hexagonal box approximately  $96 \times 96 \times 84 \text{ \AA}^3$  in volume for a total of 63,000 atoms. Electroneutrality was enforced with the addition of 150 mM NaCl.

Simulations were carried out with the CMAP corrected<sup>27</sup> CHARMM27 parameter set and the TIP3P water model. VMD<sup>28</sup> and MATLAB were used for visualization and analysis. The system was minimized using conjugate gradient minimization and heated to 310 K using Langevin dynamics with a  $10\text{-ps}^{-1}$  damping coefficient. An initial 300-ps equilibration using the NVT ensemble was carried out in which water, galactose,  $\text{Na}^+$  and heavy backbone and side-chain atoms were constrained in a harmonic potential with a force constant of  $k = 10.0 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ . We then switched to an NPT ensemble, and the restraints on the water molecules and heavy side-chain atoms were gradually removed in five steps over 1.5 ns. All remaining restraints were removed in six steps over the next 1.8 ns. Finally, 10 ns of restraint-free simulation was run. All production runs start from this equilibrated system. A Langevin piston with a 200-fs period and 100-fs decay was used to set the pressure to 1 atm. Hydrogen bond lengths were constrained with SHAKE<sup>29</sup>, and a 2-fs time step was used. A  $10 \text{ \AA}$  van der Waals cut-off was used along with the particle mesh Ewald method for the electrostatics.

Simulations with the restrained ion were carried out by holding the  $\text{Na}^+$  in a weak spherical harmonic potential using the distance from the  $\text{Na}^+$  to the centre of mass (COM) of the five coordinating residues. The equilibrium distance, based on the inward-occluded structure, was  $1.39 \text{ \AA}$ , and a force constant of  $k = 0.5 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$  produced minimal distortions in the protein. A 10-ns equilibration was run before a 200-ns production run. The r.m.s.d. of the unrestrained simulation was  $2.5 \text{ \AA}$  for the entire protein and  $3.0 \text{ \AA}$  for the restrained- $\text{Na}^+$  simulation.

**Potential of mean force calculation.** The potential of mean force (PMF) was calculated using umbrella sampling with WHAM<sup>21</sup>. We extracted 69 snapshots along the pathway and held each configuration in a harmonic potential ( $k = 7.0 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ ) with a resting length equal to the  $z$  component of the distance between the galactose COM and the binding-site COM defined by the binding-site residues. Two nanosecond trajectories were run for each umbrella, and the last 1,800 ps were used for calculating the PMF. Splitting the trajectories into two equal parts (200–1,000 ps and 1,200–2,000 ps) and computing separate PMFs revealed that the total PMF is well converged.

**Protein purification.** vSGLT proteins were cloned, expressed and purified as previously described<sup>6,30–32</sup>. Briefly, the plasmids were transformed into the TOP10 cell line expressed to OD<sub>600</sub> 1.8 and induced with 0.66 mM L-arabinose for 4 h at 29 °C. Cell membranes were isolated, solubilized with 2% decyl- $\beta$ -D-maltopyranoside and affinity-purified on a Ni-NTA column. The sample was further purified by size-exclusion chromatography (Superdex 200) and washed with crystal buffer (20 mM Tris (pH 7.5), 25 mM NaCl, 0.174% decyl- $\beta$ -D-maltopyranoside) in a 50-kDa Amicon filter unit.

**Transport assays.** Mutants were created with the QuikChange method and purified as above. vSGLT protein was reconstituted in 150 mM KCl, 10 mM Tris/Hepes (pH 8.0), 1 mM DTT, 1 mM  $\text{Na}_2\text{EDTA}$ , 1 mM  $\text{CaCl}_2$ , 1 mM  $\text{MgCl}_2$  and 0.5% decyl- $\beta$ -D-maltopyranoside, with  $1.2 \text{ mg ml}^{-1}$  sonicated lipid (90 mg asolectin soy lecithin, 10 mg cholesterol) at a protein/lipid ratio of 1:200. Addition of  $5 \text{ mg ml}^{-1}$  SM-2 Bio-Beads initiated the reconstitution and the mixture was incubated overnight at 4 °C. The proteoliposomes were collected and washed twice by centrifugation. Pelleted proteoliposomes were resuspended and underwent three freeze-thaw cycles in liquid nitrogen.

Uptake of D-galactose (88  $\mu\text{M}$ ) with  $^{14}\text{C}$ -D-galactose tracer into proteoliposomes was measured for 18 min at 22 °C in the presence or absence of 100 mM  $\text{Na}^+$  ( $\text{K}^+$  replacing  $\text{Na}^+$ ) as described previously<sup>6,30</sup>. Proteoliposomes were collected by filtration through 0.45- $\mu\text{m}$  Millipore filters and the uptake was quantified by scintillation counting. Results are expressed as the mean  $\pm$  s.e.m. of three determinations and three trials.

**Crystallization.** Protein was concentrated to  $\sim 13 \text{ mg ml}^{-1}$  before plating. Optimization by additive screening gave the best diffracting crystals with a reservoir solution containing 0.1 M MES (pH 6.5), 4% MPD and 9–13% PEG400, and tridecyl- $\beta$ -D-maltopyranoside to a final concentration of 0.0017% as an additive.

Before freezing, crystals were cryoprotected using a solution containing 30% PEG400 and 0.174% decyl- $\beta$ -D-maltopyranoside.

**Data processing, phasing and refinement.** Data was collected at 1.0  $\text{\AA}$  on cryo-cooled crystal (100 K) at the Advanced Light Source (beamline 5.0.2). Five data sets were integrated using HKL2000<sup>33</sup> and merged and subjected to B-factor-sharpening using an anisotropy correction server<sup>34</sup> (resolution cut-offs:  $a = 3.1 \text{ \AA}$ ,  $b = 2.7 \text{ \AA}$  and  $c = 2.8 \text{ \AA}$ ). Phases were calculated by molecular replacement (PHASER<sup>35</sup>) using the original vSGLT structure as a search model. The model was built in COOT<sup>36</sup> and refined using PHENIX<sup>37</sup> and BUSTER<sup>38</sup> using non-crystallographic symmetry (NCS) and TLS refinement restraints. There are two molecules per asymmetric unit with the A molecule displaying sharper electron density and lower B factors ( $88.5 \text{ \AA}^2$ ) than the B molecule ( $131.3 \text{ \AA}^2$ ). The model was built and refined to an  $R_{\text{work}}/R_{\text{free}}$  value of 25.1/27.4. The Ramachandran statistics shown areas follows: 95.5% of the residues lie in the preferred region, 4.3% lie in the allowed region and 0.2% are outliers.

The  $2F_o - F_c$  maps contained three elongated features having a maximal peak height of  $3\sigma$ . These attributes were interpreted and assigned as PEG molecules. Two are located at the periphery, whereas the third is near the Na2 site as observed in the Mhp1 structure<sup>17</sup> and is proposed to stabilize the inward-facing conformation.

The Lys 294 Ala protein crystals diffract to higher resolution than the wild-type crystals. Data from four wild-type crystals were collected and merged to achieve a 3.7  $\text{\AA}$  resolution data set. Difference Fourier maps were calculated against the final Lys 294 Ala mutant model and no significant peaks were observed. The Lys 294 Ala model was further refined using PHENIX to yield an  $R_{\text{work}}/R_{\text{free}}$  value of 30.7/34.8.

We note that although refinement was carried out with data subject to anisotropic correction, as described above, the deposited data has not been treated. Figures were created from the A-chain protomer using PYMOL<sup>39</sup>.

**Structural comparison of vSGLT with Mhp1.** Superpositions of the inward-occluded (Protein Data Bank ID, 3DH4) and inward-open conformations of vSGLT with the inward-facing conformation of Mhp1 (Protein Data Bank ID, 2X79) reveal they all share a similar global fold. The largest differences are centred near the substrate- and ion-binding sites. The Na2-site helices (TM1, TM5 and TM8) of the inward-open conformation of vSGLT have a closer fit to Mhp1 (r.m.s.d., 2.2  $\text{\AA}$ ) than the inward-occluded conformation (r.m.s.d., 2.6  $\text{\AA}$ ); thus, the inward-open vSGLT structure more closely resembles the structure of Mhp1.

26. Sali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815 (1993).
27. MacKerell, A. D. Jr, Feig, M. & Brooks, C. L. III. Improved treatment of the protein backbone in empirical force fields. *J. Am. Chem. Soc.* **126**, 698–699 (2004).
28. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
29. Ryckaert, J. P., Cicotti, G. & Berendsen, H. J. C. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of  $n$ -alkanes. *J. Comput. Phys.* **23**, 327–341 (1977).
30. Turk, E. *et al.* Molecular characterization of Vibrio parahaemolyticus vSGLT: a model for sodium-coupled sugar cotransporters. *J. Biol. Chem.* **275**, 25711–25716 (2000).
31. Turk, E., Gasyimov, O. K., Lanza, S., Horwitz, J. & Wright, E. M. A reinvestigation of the secondary structure of functionally active vSGLT, the Vibrio sodium/galactose cotransporter. *Biochemistry* **45**, 1470–1479 (2006).
32. Veenstra, M., Lanza, S., Hirayama, B. A., Turk, E. & Wright, E. M. Local conformational changes in the Vibrio  $\text{Na}^+$ /galactose cotransporter. *Biochemistry* **43**, 3620–3627 (2004).
33. Otwinowski, Z. & Minor, W. in *Macromolecular Crystallography* (eds Abelson, J. N., Simon, M. I., Carter, C. W. Jr & Sweet, R. M.) 307–326 (Methods in Enzymology 276, Academic, 1997).
34. Strong, M. *et al.* Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from Mycobacterium tuberculosis. *Proc. Natl Acad. Sci. USA* **103**, 8060–8065 (2006).
35. Qian, B. *et al.* High-resolution structure prediction and the crystallographic phase problem. *Nature* **450**, 259–264 (2007).
36. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
37. Adams, P. D. *et al.* Recent developments in the PHENIX software for automated crystallographic structure determination. *J. Synchrotron Radiat.* **11**, 53–55 (2004).
38. Joosten, R. P., Womack, T., Vriend, G. & Bricogne, G. Re-refinement from deposited X-ray data can deliver improved models for most PDB entries. *Acta Crystallogr. D* **65**, 176–185 (2009).
39. DeLano, W. L. PyMOL Molecular Viewer (<http://www.pymol.org>) (2002).



# CAREERS

**TURNING POINT** Award recognizes big discovery by young ecologist **p.994**

**ASK THE EXPERT** Advice on restarting a science career **go.nature.com/wgcuky**

**NATUREJOBS** For the latest career listings and advice **www.naturejobs.com**



LABROOTS/BIOCONFERENCE LIVE

Networking websites offer an increasing number of features, including virtual conferences, to entice researchers.

## COLLABORATION

# Social networking seeks critical mass

*There are myriad social and professional networking options for scientists. But, so far, none has proved infectious enough to go viral.*

BY VIRGINIA GEWIN

Cloning Facebook's phenomenal success has not proved easy for scientists eager to develop a network focused on shared research interests. The number of scientific networking websites is growing, but none has emerged as a 'go-to' system.

The indifference could stem from several factors: lack of confidence in individual networks; concerns that personal data might be sold on; or the fact that no one site provides tools or features valuable enough to lure a majority of busy scientists. Despite the proliferation of networks, a big question remains, says Laura James, chief operating officer at the Centre for Applied Research in Educational Technologies at the University of Cambridge, UK. What can they offer that more established sites, such as Facebook, don't?

A number of science networks are attempting to answer that question by offering extra features. ResearchGATE, based in Cambridge, Massachusetts, has launched private fee-based accounts to give institutions or companies their own sub-communities. "Activity levels in the sub-communities are much higher" than in the network at large, says Ijad Madisch, chief

executive and co-founder of the website. He says that companies want a communication solution — and ResearchGATE is happy to provide it. According to Madisch, the website has created 30 sub-communities so far, and has a total of 700,000 members around the world, the most yet achieved by any science network, he says. At its launch in May 2008, ResearchGATE, like most such sites, focused on attracting single users — but in future, it will also cater to the needs of sub-communities, says Madisch.

Some scientists may be wary of sharing too much, which probably hinders network adoption. "There is tension between collaboration and competition inherent to science," says James. "How willing will users be to share advantageous information, even with their closest collaborators, when they are still competing for grants?" Users may also be suspicious of a site's commercial intentions. "How these networks are monetized, and how their user communities will feel about the business models that emerge, remains to be seen," says James. The creators of Mendeley, a social network that also hopes to attract institutions and companies, have vowed not to sell identifiable personal information.

Aspects of some social sites have spiked in

popularity — notably, virtual events. Greg Cruikshank, chief executive of the networking site LabRoots, based in Yorba Linda, California, says that his company started BioConference Live, a series of online-only events, last year with much success. The first show, in November 2009, had more than 10,000 attendees, who were given access to 70 sessions at which they could question speakers (in future, Cruikshank expects closer to 40). They could also browse a virtual lobby, exhibit hall and exhibitor booths. "Although LabRoots does a bit of everything — from news feeds to events directory, jobs board and blogs — what distinguishes us is the virtual-event arm of the business, and we'll continue to push that," says Cruikshank.

## TOP-DOWN OR BOTTOM-UP?

The bells and whistles vary, but all the sites have the same aim: enhancing collaboration. To that end, an open-source networking movement is afoot. Rather than creating buzz among users and building a community from the ground up, the creators of freely available research-network software — such as VIVO, developed at Cornell University in Ithaca, New York, with a US\$12.2-million grant from the National Institutes of

► Health, and Harvard Catalyst Profiles, run by the Harvard Clinical and Translational Science Center in Boston — are among those working to forge a network of scientists by wooing whole institutions to adopt their platforms and upload profiles of their faculty members. “Openness is a big deal as VIVO carves out its niche,” says Mike Conlon, the site’s director.

This approach is gaining traction. On 9 October, the US Department of Agriculture (USDA) was the first federal agency to join the 14 academic institutions worldwide that contribute to VIVO’s 21,000-member network. Harvard Catalyst Profiles has more than 30,000 users, from institutions including Harvard University and the University of California, San Francisco. “We believe that data used in networks need provenance and structure — some kind of hosting institution or scientific society verifying the credibility of data,” says Conlon. Sharon Drumm, a staff officer for the USDA Agricultural Research Service in Beltsville, Maryland, says that the USDA wants VIVO to connect its 3,500 researchers with the hope of fostering collaborations, providing a single information repository for the public and allowing prospective retirees to pass on institutional knowledge.

Conlon cautions that some networking sites simply ingest publicly available data, creating profiles of people who aren’t users. Sites with no checks on data run the risk, he says, of offering less valuable information. James suggests that a high enough proportion of fake profiles could undermine entire sites. In principle, the VIVO and Profiles models are more reliable, because they use vetted data.

VIVO hopes to link profiles with not only publications or grants, but also the most important thing that scientists share — data. “Right now data are the stepchildren of this thing,” says Conlon. Linking data sets to people would allow potential collaborators to see what others can offer, he says.

But the existence of a network of profiles, even a vetted one, doesn’t necessarily mean that researchers will use it as more than just another database. “I think there’s a difference between scientists actively using a network and simply signing up all the scientists in an institution,” says Mendeley chief executive Victor Henning.

It is clear that no single site will meet all scientists’ needs. What isn’t clear is which combination of sites will. “Scientists are not really interested in social networking as an end in itself,” says Henning. “They network to boost productivity.” ■

**Virginia Gewin** is a freelance writer in Portland, Oregon.

## TURNING POINT Francesca Malfatti

*Francesca Malfatti, a postdoctoral researcher at Scripps Institution of Oceanography in San Diego, California, received the International Recognition of Professional Excellence prize in August from the International Ecology Institute in Oldendorf, Germany. The award honours young ecologists who make breakthroughs. Malfatti explains how she and her mentor found an unexpected relationship between microbes.*

### How did you get a post in the United States?

I was born in Trieste, Italy, and spent my summers collecting shells along the Adriatic Sea. I decided to study marine biology at the University of Trieste. I got a six-month grant from the Italian embassy in the United States, and could take that money anywhere in the United States. My mentors at Trieste collaborate with Farooq Azam, a marine ecologist at Scripps who specializes in microbial food webs, and he agreed to host me. We established a rapport and he has kept me on through my PhD and postdoc studies.

### How did you make a breakthrough discovery?

My research advanced significantly when Dr Azam received money from the Gordon and Betty Moore Foundation in Palo Alto, California, to buy a US\$250,000 atomic-force microscope. Using this tool, we were able to conduct our investigations at different resolutions, and discovered an unexpected relationship between bacteria that were thought to be complementary, yet isolated from one another.

### Was it difficult to confirm your findings?

It took about a year and a half to understand how best to use the atomic-force microscope. Once I mastered it, I found co-occurrence of heterotrophic bacteria and cyanobacteria. It took another three months of more-quantitative measurements to validate the discovery.

### Why is this an important discovery?

It changes how we perceive bacterial life in the open ocean. Cyanobacteria are primary producers, and get energy from the Sun; heterotrophic bacteria are the major consumers of cyanobacteria. Our hypothesis is that cyanobacteria photosynthesize part of the organic matter taken up by heterotrophs in a more tightly coupled, possibly symbiotic, system than we realized. This has implications for carbon cycling in the ocean.

### Why did you apply for the award?

Dr Azam nominated me. In his 30 years of studying marine microbial systems, he had never thought that heterotrophs and cyanobacteria



SCRIPPS INST. OCEANOGRAPHY

would be associated for any part of their lifetimes. He thought finding such an unexpected relationship was worthy of the award.

### Why are you moving on from working with him?

We will continue until I leave Scripps at the end of 2011. We have a solid working relationship, but we decided to diversify. He thinks it is best for my career and education that I see another lab or institution. So far, we have written down 21 wide-ranging ideas that we both think should be tackled.

### How will you decide how to divide the work?

Luckily, there is so much to do that it will be easy to avoid overlap or competition. We are trying to be pragmatic about who can do what type of research. For example, if I end up doing another postdoc or become a starting professor without much start-up money, I probably won’t be able to do certain kinds of experiments, such as those that require atomic-force microscopy.

### What sort of impact do you hope your research will have?

Dr Azam impressed upon me the need to create better models of dynamic ecological systems to serve policy-makers and humanity as a whole. Understanding the ocean helps humans because so many biogeochemical processes occur there. I hope my efforts help to build up an accurate ocean-based model of the planet that might help society both predict and safeguard against any impacts of climate change.

### Do you feel pressure to follow up with another breakthrough?

No. My brain doesn’t work that way. We already live in a very competitive environment. I am a self-motivator, and doing science well is good enough. Not everybody can win a Nobel Prize, but we still need strong science. ■

INTERVIEW BY VIRGINIA GEWIN



► Health, and Harvard Catalyst Profiles, run by the Harvard Clinical and Translational Science Center in Boston — are among those working to forge a network of scientists by wooing whole institutions to adopt their platforms and upload profiles of their faculty members. “Openness is a big deal as VIVO carves out its niche,” says Mike Conlon, the site’s director.

This approach is gaining traction. On 9 October, the US Department of Agriculture (USDA) was the first federal agency to join the 14 academic institutions worldwide that contribute to VIVO’s 21,000-member network. Harvard Catalyst Profiles has more than 30,000 users, from institutions including Harvard University and the University of California, San Francisco. “We believe that data used in networks need provenance and structure — some kind of hosting institution or scientific society verifying the credibility of data,” says Conlon. Sharon Drumm, a staff officer for the USDA Agricultural Research Service in Beltsville, Maryland, says that the USDA wants VIVO to connect its 3,500 researchers with the hope of fostering collaborations, providing a single information repository for the public and allowing prospective retirees to pass on institutional knowledge.

Conlon cautions that some networking sites simply ingest publicly available data, creating profiles of people who aren’t users. Sites with no checks on data run the risk, he says, of offering less valuable information. James suggests that a high enough proportion of fake profiles could undermine entire sites. In principle, the VIVO and Profiles models are more reliable, because they use vetted data.

VIVO hopes to link profiles with not only publications or grants, but also the most important thing that scientists share — data. “Right now data are the stepchildren of this thing,” says Conlon. Linking data sets to people would allow potential collaborators to see what others can offer, he says.

But the existence of a network of profiles, even a vetted one, doesn’t necessarily mean that researchers will use it as more than just another database. “I think there’s a difference between scientists actively using a network and simply signing up all the scientists in an institution,” says Mendeley chief executive Victor Henning.

It is clear that no single site will meet all scientists’ needs. What isn’t clear is which combination of sites will. “Scientists are not really interested in social networking as an end in itself,” says Henning. “They network to boost productivity.” ■

**Virginia Gewin** is a freelance writer in Portland, Oregon.

## TURNING POINT Francesca Malfatti

*Francesca Malfatti, a postdoctoral researcher at Scripps Institution of Oceanography in San Diego, California, received the International Recognition of Professional Excellence prize in August from the International Ecology Institute in Oldendorf, Germany. The award honours young ecologists who make breakthroughs. Malfatti explains how she and her mentor found an unexpected relationship between microbes.*

### How did you get a post in the United States?

I was born in Trieste, Italy, and spent my summers collecting shells along the Adriatic Sea. I decided to study marine biology at the University of Trieste. I got a six-month grant from the Italian embassy in the United States, and could take that money anywhere in the United States. My mentors at Trieste collaborate with Farooq Azam, a marine ecologist at Scripps who specializes in microbial food webs, and he agreed to host me. We established a rapport and he has kept me on through my PhD and postdoc studies.

### How did you make a breakthrough discovery?

My research advanced significantly when Dr Azam received money from the Gordon and Betty Moore Foundation in Palo Alto, California, to buy a US\$250,000 atomic-force microscope. Using this tool, we were able to conduct our investigations at different resolutions, and discovered an unexpected relationship between bacteria that were thought to be complementary, yet isolated from one another.

### Was it difficult to confirm your findings?

It took about a year and a half to understand how best to use the atomic-force microscope. Once I mastered it, I found co-occurrence of heterotrophic bacteria and cyanobacteria. It took another three months of more-quantitative measurements to validate the discovery.

### Why is this an important discovery?

It changes how we perceive bacterial life in the open ocean. Cyanobacteria are primary producers, and get energy from the Sun; heterotrophic bacteria are the major consumers of cyanobacteria. Our hypothesis is that cyanobacteria photosynthesize part of the organic matter taken up by heterotrophs in a more tightly coupled, possibly symbiotic, system than we realized. This has implications for carbon cycling in the ocean.

### Why did you apply for the award?

Dr Azam nominated me. In his 30 years of studying marine microbial systems, he had never thought that heterotrophs and cyanobacteria



SCRIPPS INST. OCEANOGRAPHY

would be associated for any part of their lifetimes. He thought finding such an unexpected relationship was worthy of the award.

### Why are you moving on from working with him?

We will continue until I leave Scripps at the end of 2011. We have a solid working relationship, but we decided to diversify. He thinks it is best for my career and education that I see another lab or institution. So far, we have written down 21 wide-ranging ideas that we both think should be tackled.

### How will you decide how to divide the work?

Luckily, there is so much to do that it will be easy to avoid overlap or competition. We are trying to be pragmatic about who can do what type of research. For example, if I end up doing another postdoc or become a starting professor without much start-up money, I probably won’t be able to do certain kinds of experiments, such as those that require atomic-force microscopy.

### What sort of impact do you hope your research will have?

Dr Azam impressed upon me the need to create better models of dynamic ecological systems to serve policy-makers and humanity as a whole. Understanding the ocean helps humans because so many biogeochemical processes occur there. I hope my efforts help to build up an accurate ocean-based model of the planet that might help society both predict and safeguard against any impacts of climate change.

### Do you feel pressure to follow up with another breakthrough?

No. My brain doesn’t work that way. We already live in a very competitive environment. I am a self-motivator, and doing science well is good enough. Not everybody can win a Nobel Prize, but we still need strong science. ■

INTERVIEW BY VIRGINIA GEWIN



# SILENCE

*Turn on, tune out.*

BY JOHN FRIZELL

“**Y**ou have to do something about this awful music,” said Ellie.  
“I don’t play it loud.”

“I didn’t mean you. It’s ... everyone.”

“I can’t be responsible for the musical tastes of the entire world.”

A deliberate misunderstanding was typical of her little brother. She glared at him.

“OK. I’ll get some stuff.”

They were meeting in the neutral ground of the living room. Jamie sloped off to the seething maelstrom of mechanical creatures and things he called a room and returned carrying a red plastic box. She looked into it as he put it down. Nothing was moving.

Jamie put *Out of Space* on the music system and turned it up until she winced.

“That’s ghastly.”

“How’s this?”

He fished something that looked like a big set of orange earmuffs out of the box and she put them on. The dreadful beat softened, almost vanished.

“That’s better, but ...”

Her voice sounded funny. It wasn’t just the music, she couldn’t hear herself, couldn’t hear anything.

“What are these?” she said, handing them back.

“Ear defenders. I wear them for grinding or cutting or when you and Mum argue.”

“It’s just the music I don’t want to hear. I have to be able to hear everything else.”

“That’s going to cost you.”

They settled on a month of gourmet desserts, prepared by Ellie.

“Starting tomorrow night.”

“Starting when it works.”

“It will be working tomorrow night.”

Ellie put on the headphones Jamie handed her and looked dubiously at the pile of circuit boards into which they were plugged. She could see bare strips of copper on two of the boards. It didn’t look safe.

“Test version. Final result will be wireless and fit in your purse. Check it out.”

He cranked up *Out of Space*. The beat pounded on her ears.

“This doesn’t work. I can still ...”

Abruptly the relentless beat died away.

“What do you think?”

She could hear him clearly, but not the music. She lifted one side of the headphones to check it was still on and quickly clamped them back.

“It didn’t work at first but now it’s perfect.”

Jamie smiled.

“It samples the track then it phones a service to get the name. That’s why you can hear the music for a few seconds. Once it knows the track it selects it from its database, syncs it up and then uses the waveform of the track to block out the incoming waveform.”

She was not quite sure what he meant but could see the weaknesses.

“So I’m always going to get a blast of sound, and if someone plays a song that isn’t on the database, I’m going to hear it.”

“Well, yes.”

“It’s really good,” she said, using her best positive, reinforcing voice, “but not quite what I need. I’ll do you *crème brûlée* tonight and it won’t count as part of your month’s total.”

Jamie could make anything work and, with

wearing headphones all day.”

“Why not?”

“Because I’d look like an idiot,” occurred to her but Jamie didn’t understand anything to do with look or style.

“Some kids wear headphones all the time they aren’t in class. Girls even.”

“Losers,” thought Ellie. Aloud she said: “I wouldn’t want the people whose music I’m tuning out to know.”

He looked at her thoughtfully.

“I like the way you wear your hair.”

Amazing! A compliment from Jamie. He was actually noticing her look. If he cared about how other people looked she could get him to care about how he looked. And once that was accomplished she could find him a girlfriend. Her mind raced, sorting though the younger sisters of her friends.

“I’ll get some Eartalkers or hearing aid



her encouragement, he would. She went to the kitchen, got a vanilla pod from her spice drawer, split it lengthwise, scraped the seeds out while thinking of the social advantages of actually being able to follow conversations at parties, and then she cut the pod up into small pieces and cracked six eggs, separating each into white and yolk. She would find something to do with the whites later.

It took Jamie more than a week but the results were spectacular. Admittedly she had to train the system by pressing a button when she heard music she didn’t like, but once she had done it the music never came back, not even if she pressed the training button halfway through and the music restarted from the beginning. Jamie went on about tonal analysis, pattern recognition, neural net processing, context algorithms and other mysterious stuff. It sounded like he would get another patent out of it and even more money in his trust fund.

“Just one little thing. These headphones are nice and very comfortable but I can’t go around

inserts that are retained by your ears. The hair will cover them. Stealth.”

“Oh. Well. See you in the kitchen.”

Jamie loved rich desserts — nothing stuck to his skinny frame. He was shovelling down the syllabub. Ellie allowed herself only a single teaspoon, just to get the taste. She had zested a lime and a lemon for this and she was convinced that the combined citrus taste was much better than the standard recipe of lemon alone. She had added a bit of cardamom as well, giving the traditional English dessert a hint of something Middle Eastern and exotic. She decided on a second teaspoon.

Jamie was talking happily between bites about how he had accomplished this latest feat of engineering. She smiled and nodded in the pauses. The thing could detect more patterns than just the ones in music — she could see his lips moving but she could not hear a word he said. ■

*John Frizell was trained in biochemistry and works on ocean conservation for Greenpeace. In his spare time he walks, builds robots and writes short stories.*

➤ **NATURE.COM**  
Follow Futures on  
Facebook at:  
[go.nature.com/mtoodm](http://go.nature.com/mtoodm)

# Was the universal common ancestry proved?

ARISING FROM D. L. Theobald *Nature* **465**, 219–222 (2010)

The question of whether or not all life on Earth shares a single common ancestor has been a central problem of evolutionary biology since Darwin<sup>1</sup>. Although the theory of universal common ancestry (UCA) has gathered a compelling list of circumstantial evidence, as given in ref. 2, there has been no attempt to test statistically the UCA hypothesis among the three domains of life (eubacteria, archaeobacteria and eukaryotes) by using molecular sequences. Theobald<sup>2</sup> recently challenged this problem with a formal statistical test, and concluded that the UCA hypothesis holds. Although his attempt is the first step towards establishing the UCA theory with a solid statistical basis, we think that the test of Theobald<sup>2</sup> is not sufficient enough to reject the alternative hypothesis of the separate origins of life, despite the Akaike information criterion (AIC) of model selection<sup>3</sup> giving a clear distinction between the competing hypotheses.

Dawkins<sup>4</sup> argued that even though it may, at first, seem unlikely that such a complex structure as the eye evolved by selection, it could have been realized by a long sequence of small evolutionary steps driven by selection. Theobald<sup>2</sup> mentions that statistically significant sequence similarity can arise from factors other than common ancestry, such as convergent evolution due to selection, but such factors were not taken into account in his ‘formal’ test to reject the independent origins hypothesis.

Table 1 shows that the formal test provides support for a common origin of two putatively unrelated genes, mitochondrial *cytb* and *nd2*, with no homology. However, we believe that this result should not be regarded as evidence of the ultimate common ancestry of *cytb* and *nd2*. This raises a question mark as to the effectiveness of the formal

test applied by Theobald<sup>2</sup>. It should be noted that, because alignment gives a bias for common ancestry, we did not make an alignment between *cytb* and *nd2*. To reject the separate origins hypothesis of the domains of life, it would be indispensable to develop a more ‘biological’ test to show that even by improving the model of the separate origins by taking into account biological factors such as the possibility of convergent evolution due to selection, the UCA hypothesis is still supported by the AIC. To do this, it is necessary to develop an entirely new methodological framework of molecular phylogenetics that is different from the conventional framework that neglects convergent and parallel evolution. Notably, there have been many reported cases of convergent and parallel evolution misleading molecular phylogenetic inference<sup>5–9</sup>, and such a method is needed for molecular phylogenetics in general.

Takahiro Yonezawa<sup>1</sup> & Masami Hasegawa<sup>1,2</sup>

<sup>1</sup>School of Life Sciences, Fudan University, Shanghai 200433, China.

e-mail: masamihase@gmail.com

<sup>2</sup>Institute of Statistical Mathematics, Tokyo 190-8562, Japan.

Received 2 July; accepted 24 August 2010.

**Table 1 | A formal test of the common ancestry between mitochondrial genes *cytb* and *nd2***

Test statistic	Score or value	Number of parameters
Common origin		
InL ( <i>cytb</i> + <i>nd2</i> )	–5,090.20	18
AIC	10,216.4	
Independent origin		
InL ( <i>cytb</i> )	–2,503.82	12
InL ( <i>nd2</i> )	–2,608.17	12
Total InL	–5,111.99	24
AIC	10,271.97	

Nucleotide sequences of the mitochondrial genes *cytb* and *nd2* from cow, deer and hippopotamus were analysed by PAML<sup>10</sup> with the GTR +  $\Gamma$  model assuming the relations of ((cow, deer), hippopotamus) for the common origin model. The 5'-terminal 1,038 bp (excluding the initiation codon) were used without making further alignment between the two different genes. The common origin model gave a lower AIC value than the independent origin model. InL, log-likelihood score.

1. Darwin, C. *On the Origin of Species by Means of Natural Selection, or, The Preservation of Favoured Races in the Struggle for Life* (J. Murray, 1859).
2. Theobald, D. L. A formal test of the theory of universal common ancestry. *Nature* **465**, 219–222 (2010).
3. Akaike, H. in *Second International Symposium on Information Theory* (eds Petrov, B. N. & Csaki, F.) 267–281 (Akademai Kiado, 1973).
4. Dawkins, R. *The Blind Watchmaker: Why the Evidence of Evolution Reveals a Universe without Design* (Longman Scientific & Technical, 1986).
5. Stewart, C.-B., Schilling, J. W. & Wilson, A. C. Adaptive evolution in the stomach lysozymes of foregut fermenters. *Nature* **330**, 401–404 (1987).
6. Christin, P. A., Salamin, N., Savolainen, V., Duvall, M. R. & Besnard, G. C<sub>4</sub> photosynthesis evolved in grasses via parallel adaptive genetic changes. *Curr. Biol.* **17**, 1241–1247 (2007).
7. Liu, Y. *et al.* Convergent sequence evolution between echolocating bats and dolphins. *Curr. Biol.* **20**, R53–R54 (2010).
8. Li, Y., Liu, Z., Shi, P. & Zhang, J. The hearing gene Prestin unites echolocating bats and whales. *Curr. Biol.* **20**, R55–R56 (2010).
9. Zhong, B., Yonezawa, T., Zhong, Y. & Hasegawa, M. The position of Gnetales among seed plants: overcoming pitfalls of chloroplast phylogenomics. *Mol. Biol. Evol.* doi:10.1093/molbev/msq170 (2 July 2010).
10. Yang, Z. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).

**Author Contributions** T.Y. and M.H. designed the work, T.Y. carried out the analysis, and M.H. wrote the manuscript.

**Competing financial interests:** declared none.

doi:10.1038/nature09482

## Theobald reply

REPLYING TO T. Yonezawa & M. Hasegawa *Nature* **468**, doi:10.1038/nature09482 (2010)

Yonezawa and Hasegawa<sup>1</sup> provide an example from two apparently unrelated families of nucleic acid coding sequences for which an Akaike information criterion (AIC) model selection test, similar to mine<sup>2</sup>, chooses a common origin hypothesis. Although this may seem surprising, the coding sequences in this example were aligned in the same reading frame. The constraints of the genetic code are expected to induce correlations between these sequences (and among all coding sequences) that are not due to common ancestry. For instance, owing to codon bias and the structure of the genetic code, in these sequences the second codon position is biased towards T (about twofold over average), whereas the third position is usually an A (~50%) and rarely a G (~4%).

One can account for these correlations explicitly by using codon models (as implemented in PAML<sup>3</sup>, codonFreq = 2 or 3) or standard amino acid models (as in PhyML<sup>4</sup>). With these more realistic models, independent ancestry is the strongly preferred hypothesis. Furthermore, the raw likelihoods and AIC scores increase significantly (by hundreds to thousands of logs), indicating that codon and amino acid models are greatly superior to the naive nucleotide models.

Yonezawa and Hasegawa<sup>1</sup> point out that I<sup>2</sup> did not explicitly test models in which selection or biophysical constraints generate sequence correlations among proteins with independent origins. Formal phylogenetic models accounting for such factors are currently unavailable; their development would be a welcome advance. Although these are important considerations for proteins with low sequence similarity, neither selection nor physical constraints alone can plausibly generate the high levels of sequence similarity (>55% average sequence identity) observed in the universal protein data set that I used<sup>2,5</sup>. The amount of adaptive convergence necessary to produce thousands of identical amino acids among 23 different proteins from completely independent beginnings is not comparable to the limited molecular convergence seen with, for example, homologous digestive lysozymes<sup>6</sup>, in which already highly similar proteins (in function, structure and sequence) later acquired a handful of identical substitutions in parallel.

How could selection or biophysical constraints induce correlations among unrelated sequences? If certain similar amino acid sequences are necessary for performing specific functions (or for adopting a specific tertiary conformation that is necessary for function), then selection for function may 'lead' proteins with independent origins to neighbouring regions of sequence space. However, no particular protein sequence or fold is necessary for any given function. There are abundant examples of proteins with undetectable sequence similarity and different folds that perform the same biochemical and cellular functions<sup>7</sup>. For example, the proteases subtilisin, trypsin and carboxypeptidase have the same active site and mechanism, whereas papain, renin and thermolysin have different active sites and different mechanisms. All six proteases have radically different folds and sequences. Because different folds in general have different sequence requirements, proteins with the same function need not have similar sequences.

Even assuming that a certain protein fold is necessary for a given function, current molecular evidence indicates that sequence requirements for a fold are extremely low—nearly indistinguishable from random. This data comes from many independent sources from throughout biology.

Many large classes of proteins with identical folds have no detectable sequence similarity (for example, families of TIM barrels, carbonic

anhydrases, OB-folds, SH3 domains, Rossmann folds and immunoglobulin domains). These proteins provide *prima facie* evidence that sequence requirements for any particular fold and function are nearly indistinguishable from random. Protein domains in the SCOP database<sup>8</sup> from different superfamilies yet with the same fold share ~9% sequence identity<sup>9</sup>.

Identical folds with known independent origins have nearly random sequence similarity<sup>9,10</sup>. For example, unrelated proteins with the same fold from the MALISAM database share  $8.5 \pm 0.4\%$  sequence identity<sup>9,10</sup>. This data can be used to estimate the correlations among independently evolved and created proteins with the same fold, and the correlations are nearly random. In the universal protein data set that I used<sup>2</sup>, the average sequence correlation induced by common ancestry is roughly one log-likelihood per site for the most divergent proteins. In contrast, the correlations among independent proteins with the same fold are ~100 times weaker. From this we can estimate that model selection scores for common ancestry hypotheses will be many thousands of logs greater than competing selection hypotheses.

Even the most conserved proteins have not yet reached the limits of sequence space, which has been estimated to be near the random expectation for any given fold and function<sup>11</sup>.

These arguments are largely circumstantial and informal. I have not tested all possible competing hypotheses, and my analysis will not be the "last word on common ancestry"<sup>12</sup>. I emphasize that I have in no sense provided an absolute 'proof' of universal common ancestry. One of the great advantages of the model selection framework that I presented is that if a novel model is proposed with a well-defined likelihood function, then we can easily compare it to the common ancestry models and see how it fares.

D. L. Theobald<sup>1</sup>

<sup>1</sup>Department of Biochemistry, Brandeis University, Waltham, Massachusetts 01778, USA.

e-mail: dtheobald@brandeis.edu

1. Yonezawa, T. & Hasegawa, M. Was the universal common ancestry proved? *Nature* **468**, 10.1038/nature09482 (2010).
2. Theobald, D. L. A formal test of the theory of universal common ancestry. *Nature* **465**, 219–222 (2010).
3. Yang, Z. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **13**, 555–556 (1997).
4. Guindon, S. & Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696–704 (2003).
5. Brown, J. R., Douady, C. J., Italia, M. J., Marshall, W. E. & Stanhope, M. J. Universal trees based on large combined protein sequence data sets. *Nature Genet.* **28**, 281–285 (2001).
6. Stewart, C. B., Schilling, J. W. & Wilson, A. C. Adaptive evolution in the stomach lysozymes of foregut fermenters. *Nature* **330**, 401–404 (1987).
7. Omelchenko, M. V., Galperin, M. Y., Wolf, Y. I. & Koonin, E. V. Non-homologous isofunctional enzymes: a systematic analysis of alternative solutions in enzyme evolution. *Biol. Direct* **5**, 31 (2010).
8. Andreeva, A. *et al.* Data growth and its impact on the SCOP database: new developments. *Nucleic Acids Res.* **36**, D419–D425 (2008).
9. Cheng, H., Kim, B. H. & Grishin, N. V. Discrimination between distant homologs and structural analogs: lessons from manually constructed, reliable data sets. *J. Mol. Biol.* **377**, 1265–1278 (2008).
10. Cheng, H., Kim, B. H. & Grishin, N. V. MALISAM: a database of structurally analogous motifs in proteins. *Nucleic Acids Res.* **36**, D211–D217 (2008).
11. Povolotskaya, I. S. & Kondrashov, F. A. Sequence space and the ongoing expansion of the protein universe. *Nature* **465**, 922–926 (2010).
12. Steel, M. & Penny, D. Origins of life: Common ancestry put to the test. *Nature* **465**, 168–169 (2010).

doi:10.1038/nature09483



## Global systematics of arc volcano position

ARISING FROM Grove, T. *et al.* *Nature* **459**, 694–697 (2009)

Global systematics in the location of volcanic arcs above subduction zones<sup>1,2</sup> are widely considered to be a clue to the melting processes that occur at depth, and the locations of the arcs have often been explained in terms of the release of hydrous fluids near the top of the subducting slab (see, for example, refs 3–6). Grove *et al.*<sup>7</sup> conclude that arc volcano location is controlled by melting in the mantle at temperatures above the water-saturated upper-mantle solidus and below the upper limit of stability of the mineral chlorite and in particular, that the arc fronts lie directly above the shallowest point of such melt regions in the mantle. Here we show that this conclusion is incorrect because the calculated arc locations of Grove *et al.*<sup>7</sup> are in error owing to the inadequate spatial resolution of their numerical models, and because the agreement that they find between predicted and observed systematics arises from a spurious correlation between calculated arc location and slab dip. A more informative conclusion to draw from their experiments is that the limits of chlorite stability (figure 1b of ref. 7) cannot explain the global systematics in the depth to the slab beneath the sharply localized arc fronts.

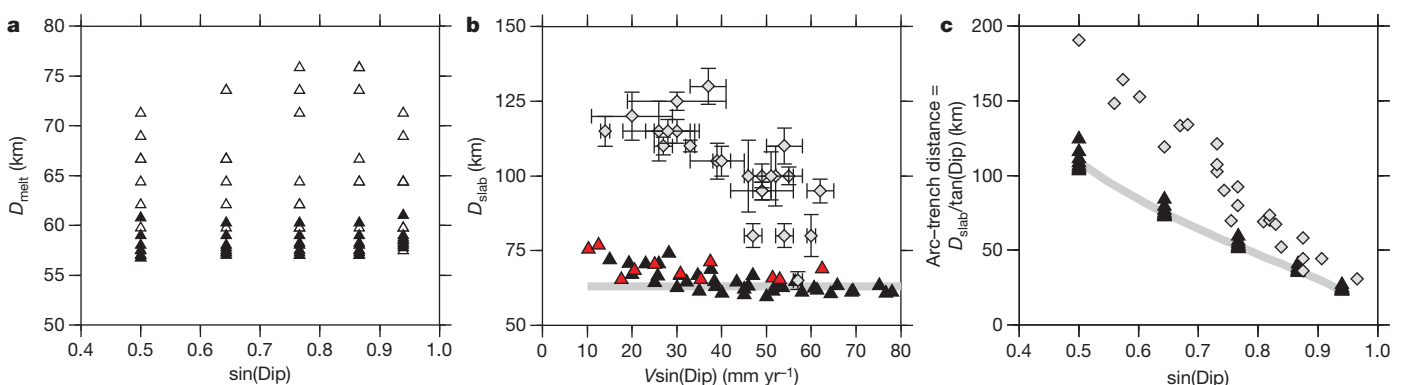
Grove *et al.*<sup>7</sup> hypothesize that arc volcano location is controlled by melting in the mantle at pressure and temperature conditions defined as ' $P, T_{\text{melt}}$ ' in their figure 1b. Grove *et al.*<sup>7</sup> then use numerical models of subduction zones to predict arc location and its global systematics. They conclude that the agreement between their calculated systematics of arc location and observations of real subduction zones<sup>2,8</sup> validates their hypothesis (figure 3 of ref. 7) but closer inspection of the shape of the  $P, T_{\text{melt}}$  region casts doubt upon this conclusion. A characteristic feature of subduction-zone models<sup>9</sup> is the narrow thermal boundary layer, sub-parallel to and just above the slab surface, which contains the temperature range of  $P, T_{\text{melt}}$  ( $\sim 800$ – $850$  °C). For all but the slowest convergence rates, this boundary layer begins close to the depth at which the slab is viscously coupled to the wedge. Hence we should expect the region enclosing  $P, T_{\text{melt}}$  to be a very thin, continuous layer above the slab, with its shallowest extent at an almost constant depth.

The results of Grove *et al.*<sup>7</sup> (green squares in their figure 2) are inconsistent with this expectation, and raise the suspicion of an error in their calculations.

To locate their region of  $P, T_{\text{melt}}$ , Grove *et al.*<sup>7</sup> determined which nodes of their  $2.3 \times 2.3$ -km computational mesh lay within that  $P$ - $T$  range. Because those conditions occur within a boundary layer only a few kilometres thick that is inclined at an angle to the mesh, this procedure did not resolve the full extent of the  $P, T_{\text{melt}}$  region. To check their results, we calculated the temperature fields for subduction zones on a  $1 \times 1$ -km grid, then resampled it to both  $2.3$ -km resolution and to  $0.25$ -km resolution. This was done for a range of subduction parameters and for each calculation we determined the  $P, T_{\text{melt}}$  region and its shallowest point. We found that at  $2.3$ -km resolution, the minimum depth of  $P, T_{\text{melt}}$  ranged between about 57 and 76 km, consistent with the range found by Grove *et al.*<sup>7</sup>. On the  $0.25 \times 0.25$ -km grid, however, the minimum depth was confined between 57 and 61 km (Fig. 1a), consistent with the expectations we describe in the preceding paragraph. At either resolution, the minimum depth of  $P, T_{\text{melt}}$  is independent of the slab dip and of the convergence rate.

Grove *et al.*<sup>7</sup> compare their calculations with seismic studies, which show that the depth of the slab beneath arcs varies between  $\sim 80$  and  $\sim 150$  km (refs 2, 8) and has a negative correlation with the descent speed of the slab (Fig. 1b). The depth to the top of the slab predicted by the hypothesis of Grove *et al.*<sup>7</sup> applied under our recalculations is  $\sim 60$ – $75$  km, independent of dip or convergence rate (Fig. 1b), and thus does not agree with the observations.

The agreement between model and observations in Grove *et al.*<sup>7</sup> is spurious, and is the result of their choice of variables. Figure 1c recreates their figure 3, which shows the apparent consistency between model and observations, using our recalculated location of arcs. The sine of slab dip is plotted on the  $x$  axis, and on the  $y$  axis is the arc-trench distance, which for all points (calculated and observed; see



**Figure 1 | Arc position versus subduction parameters for data and models.** **a**, Calculated depth  $D_{\text{melt}}$  of the shallowest portion of the  $P, T_{\text{melt}}$ -based melting field (compare figures 1 and 2 of ref. 7). Calculations were carried out on a 1-km finite-volume mesh<sup>9</sup>, for dip of  $30^\circ$  to  $70^\circ$  in steps of  $10^\circ$ , and for convergence rate  $V$  from 30 to 100 mm yr<sup>-1</sup>, in steps of 10 mm yr<sup>-1</sup>; these ranges include the parameters of the calculations of ref. 7. The points correspond to the minimum depths of melting calculated according to the hypothesis and methods of Grove *et al.*<sup>7</sup> for a  $2.3 \times 2.3$ -km resampled grid (open triangles) and for a  $0.25 \times 0.25$ -km resampled grid (filled triangles). **b**, Diamonds show the depth of the slab  $D_{\text{slab}}$ , determined seismologically<sup>2</sup> (error bars as described by ref. 2); filled

triangles show the calculated  $D_{\text{slab}}$  below the locus of shallowest melting, for the  $0.25 \times 0.25$ -km resampled grid from panel a. The red triangles correspond to the corrected values of  $D_{\text{slab}}$  for the combinations of dip and convergence rate used by ref. 7 (T. Grove *et al.*, personal communication). The grey line corresponds to a constant  $D_{\text{slab}} = 62$  km. **c**, This panel corresponds to the lower 200 km of figure 3 in ref. 7. Points as for panel b, plotted for the horizontal distance between the trench and the arc, which is equal to  $D_{\text{slab}}/\tan(\text{Dip})$ , the quantity on the  $y$  axis of figure 3 of ref. 7. The grey line corresponds to  $D_{\text{slab}} = 62$  km and demonstrates the spurious correlation referred to in the main text.

table 1 in ref. 7) is taken as the depth of the slab divided by the tangent of the dip. The presence of the sine of the dip on each axis ensures a spurious correlation; this is illustrated clearly in Fig. 1c by the grey line that corresponds to a constant value of the depth of the slab,  $D_{\text{slab}} = 62$  km.

Therefore there is no significance in the match between models and observations reported by Grove *et al.*<sup>7</sup>, and their conclusion that “the kinematic control on the location of mantle melting is primarily slab dip” (page 696 of ref. 7) is mistaken. Instead, we conclude from their experiments that the limits of chlorite stability (figure 1b of ref. 7) cannot explain the global systematics in the depth of the slab beneath sharply localized arc fronts, which is true for any strongly temperature-dependent process that takes place near the top of the slab, as we have discussed. In ref. 10 we suggest a process that can account for the global systematics in location of the arcs.

**Philip C. England<sup>1</sup> & Richard F. Katz<sup>1</sup>**

<sup>1</sup>Department of Earth Sciences, Parks Road, Oxford OX1 3PR, UK.  
e-mail: philip.england@earth.ox.ac.uk

Received 7 September 2009; accepted 12 April 2010.

1. Tovish, A. & Schubert, G. Island arc curvature, velocity of convergence and angle of subduction. *Geophys. Res. Lett.* **5**, 329–332 (1978).
2. England, P., Engdahl, R. & Thatcher, W. Systematic variation in the depths of slabs beneath arc volcanoes. *Geophys. J. Int.* **156**, 377–408 (2004).
3. Gill, J. *Orogenic Andesites and Plate Tectonics* (Springer, 1981).
4. Tatsumi, Y. & Eggins, S. *Subduction Zone Magmatism* (Blackwell Science, 1995).
5. Iwamori, H. Transportation of H<sub>2</sub>O and melting in subduction zones. *Earth Planet. Sci. Lett.* **160**, 65–80 (1998).
6. Tatsumi, Y. The subduction factory: how it operates in the evolving Earth. *GSA Today* **15**, 4–10 (2005).
7. Grove, T., Till, C., Lev, E., Chatterjee, N. & Medard, E. Kinematic variables and water transport control the formation and location of arc volcanoes. *Nature* **459**, 694–697 (2009); erratum **460**, 1044 (2009).
8. Syracuse, E. & Abers, G. Global compilation of variations in slab depth beneath arc volcanoes and implications. *Geochem. Geophys. Geosyst.* **7**, Q05017, doi:10.1029/2005GC001045 (2006).
9. van Keken, P. *et al.* A community benchmark for subduction zone modeling. *Phys. Earth Planet. Inter.* **171**, 187–197 (2008).
10. England, P. C. & Katz, R. F. Melting above the anhydrous solidus controls the location of volcanic arcs. *Nature* **467**, 700–703 (2010).

**Competing financial interests:** declared none.

doi:10.1038/nature09154

## Grove *et al.* reply

REPLYING TO England, P. C. & Katz, R. F. *Nature* **468**, doi:10.1038/nature09154 (2010)

In their Comment England and Katz<sup>1</sup> suggest that our model<sup>2</sup> contains two flaws and that there are additional problems in our thermal models. This Reply points out an important part of our model that England and Katz<sup>1</sup> appear to have missed, addresses their suggestion that there are flaws and discusses whether our thermal models are in error.

The Comment<sup>1</sup> states that we “conclude that the arc fronts lie directly above the shallowest point [that satisfies the  $P$ ,  $T_{\text{melt}}$  criterion] in the mantle”. This corresponds to our path A in figure 1 of ref. 2. The  $P$ ,  $T_{\text{melt}}$  criterion described in ref. 1 refers to the melting that initiates just above the slab over a range of depths, illustrated between paths A and C in figure 1 of ref. 2. As we discussed<sup>2</sup>, when these initial melts ascend into the overlying mantle wedge, not all of them will experience a pressure–temperature path that allows them to erupt from an arc volcano on the Earth’s surface. The melts formed at the shallowest depths (path A) will encounter cooler mantle as they ascend into the overlying mantle wedge and these melts will freeze. Only melts that ascend into the hottest interior portion of the mantle wedge (such as path B in figure 1 of ref. 2) will undergo sufficient melting to produce arc front volcanoes. To summarize our findings<sup>2</sup>, there are two important factors that control the location of arc volcanoes: (1) chlorite dehydration releases H<sub>2</sub>O near the slab–wedge interface, and the H<sub>2</sub>O ascends into overlying mantle that is above the H<sub>2</sub>O-saturated mantle solidus ( $P$ ,  $T_{\text{melt}}$  in figure 1b of ref. 2) and (2) the temperature of the overlying mantle wedge increases with decreasing pressure to allow flux melting to continue to high extents and allow these high-extent melts to erupt at arc volcanoes (path B in figure 1 of ref. 2).

England and Katz also state that the agreement that we<sup>2</sup> “find between predicted and observed systematics arises from a spurious correlation between calculated arc location and slab dip” (ref. 1). They attribute this purported spurious correlation in our figure 3 (ref. 2) to the presence of the tangent function on the vertical axis and a sine function on the horizontal axis. Although there is trigonometry involved in the correlation shown on this figure<sup>2</sup>, the relations are

not spurious and are meaningful. The salient point in our figure 3 is that the beginning of H<sub>2</sub>O-saturated melting in our modelling (path A in figure 1a of ref. 2) consistently occurs at a depth of 60–70 km near the slab–wedge interface and is independent of the convergence rate and dip. We point out that these shallowest melts do not reach the surface (figure 1 of ref. 2), nor do they influence the location of volcanoes. Instead, the maximum amount of melting and hence the location of arc volcanoes are controlled by the position of the hottest part of the wedge above a slab. This is the region between paths B and C (figure 1 of ref. 2), the region of maximum melting from our models. The arc–trench distance for paths B to C, and thus the location of arc volcanoes, is close to the values reported by England *et al.*<sup>3</sup> and parallel to the trend of Syracuse and Abers<sup>4</sup>. The distance a given isotherm is from the trench decreases with increasing convergence rate and spans a range of values that are represented in the data of ref. 3. An interesting outcome of our thermal modelling (figure 2 of ref. 2) is that at steep dip angles, paths A and B occur at very similar distances from the trench.

England and Katz say that our thermal modelling results (in figure 2 of ref. 2) “raise the suspicion of an error in [our] calculations” (ref. 1). England and Katz continue with a discussion of grid size in the numerical calculations that they performed, but it is impossible for us or for any reader of the Comment<sup>1</sup> to assess the veracity of their claim that we<sup>2</sup> “did not resolve the full extent of the  $P$ ,  $T_{\text{melt}}$  region” (ref. 1). We have verified our modelling methods using the community benchmarks developed for subduction zone modelling<sup>5</sup> and we also find that our model results for the temperature structure near the slab–wedge interface are comparable to those of others who have benchmarked their models, such as Wada and Wang<sup>6</sup>, who explicitly considered the issues associated with slab–mantle viscous coupling.

Thus, we disagree with the conclusion reached by England and Katz<sup>1</sup> that “the limits of chlorite stability cannot explain the global systematics in the depth of the slab beneath sharply localized arc fronts”. The conclusions we reached in ref. 2 rely on the interplay

of two important controls on hydrous melting in the mantle wedge above subducted slabs: the dehydration of chlorite near the base of the wedge and the temperature structure of the overlying mantle wedge.

**T. L. Grove<sup>1</sup>, C. B. Till<sup>1</sup>, E. Lev<sup>2</sup>, N. Chatterjee<sup>1</sup> & E. Médard<sup>3</sup>**

<sup>1</sup>Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

e-mail: [tlgrove@mit.edu](mailto:tlgrove@mit.edu)

<sup>2</sup>Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York 10964, USA.

<sup>3</sup>Laboratoire Magmas et Volcans, Université Blaise Pascal, Clermont-Ferrand, F-63038, France.

1. England, P. C. & Katz, R. F. Global systematics of arc volcano position. **468**, doi:10.1038/nature09154 (this issue).
2. Grove, T., Till, C., Lev, E., Chatterjee, N. & Medard, E. Kinematic variables and water transport control the formation and location of arc volcanoes. *Nature* **459**, 694–697 (2009); erratum **460**, 1044 (2009).
3. England, P., Engdahl, R. & Thatcher, W. Systematic variation in the depths of slabs beneath arc volcanoes. *Geophys. J. Int.* **156**, 377–408 (2004).
4. Syracuse, E. M. & Abers, G. A. Global compilation of variations in slab depth beneath arc volcanoes and implications. *Geochem. Geophys. Geosyst.* **7**, Q05017, doi: 10.1029/2005GC001045 (2006).
5. van Keken, P. E., *et al.* A community benchmark for subduction zone modeling. *Phys. Earth Planet. Inter.* **171**, 187–197 (2008).
6. Wada, I. & Wang, K. Common depth of slab-mantle decoupling: reconciling diversity and uniformity of subduction zones. *Geochem. Geophys. Geosyst.* **10**, Q10009, doi: 10.1029/2009GC002570 (2009).

doi:10.1038/nature09155



# Was the universal common ancestry proved?

ARISING FROM D. L. Theobald *Nature* **465**, 219–222 (2010)

The question of whether or not all life on Earth shares a single common ancestor has been a central problem of evolutionary biology since Darwin<sup>1</sup>. Although the theory of universal common ancestry (UCA) has gathered a compelling list of circumstantial evidence, as given in ref. 2, there has been no attempt to test statistically the UCA hypothesis among the three domains of life (eubacteria, archaeobacteria and eukaryotes) by using molecular sequences. Theobald<sup>2</sup> recently challenged this problem with a formal statistical test, and concluded that the UCA hypothesis holds. Although his attempt is the first step towards establishing the UCA theory with a solid statistical basis, we think that the test of Theobald<sup>2</sup> is not sufficient enough to reject the alternative hypothesis of the separate origins of life, despite the Akaike information criterion (AIC) of model selection<sup>3</sup> giving a clear distinction between the competing hypotheses.

Dawkins<sup>4</sup> argued that even though it may, at first, seem unlikely that such a complex structure as the eye evolved by selection, it could have been realized by a long sequence of small evolutionary steps driven by selection. Theobald<sup>2</sup> mentions that statistically significant sequence similarity can arise from factors other than common ancestry, such as convergent evolution due to selection, but such factors were not taken into account in his ‘formal’ test to reject the independent origins hypothesis.

Table 1 shows that the formal test provides support for a common origin of two putatively unrelated genes, mitochondrial *cytb* and *nd2*, with no homology. However, we believe that this result should not be regarded as evidence of the ultimate common ancestry of *cytb* and *nd2*. This raises a question mark as to the effectiveness of the formal

test applied by Theobald<sup>2</sup>. It should be noted that, because alignment gives a bias for common ancestry, we did not make an alignment between *cytb* and *nd2*. To reject the separate origins hypothesis of the domains of life, it would be indispensable to develop a more ‘biological’ test to show that even by improving the model of the separate origins by taking into account biological factors such as the possibility of convergent evolution due to selection, the UCA hypothesis is still supported by the AIC. To do this, it is necessary to develop an entirely new methodological framework of molecular phylogenetics that is different from the conventional framework that neglects convergent and parallel evolution. Notably, there have been many reported cases of convergent and parallel evolution misleading molecular phylogenetic inference<sup>5–9</sup>, and such a method is needed for molecular phylogenetics in general.

Takahiro Yonezawa<sup>1</sup> & Masami Hasegawa<sup>1,2</sup>

<sup>1</sup>School of Life Sciences, Fudan University, Shanghai 200433, China.

e-mail: masamihase@gmail.com

<sup>2</sup>Institute of Statistical Mathematics, Tokyo 190-8562, Japan.

Received 2 July; accepted 24 August 2010.

**Table 1 | A formal test of the common ancestry between mitochondrial genes *cytb* and *nd2***

Test statistic	Score or value	Number of parameters
Common origin		
InL ( <i>cytb</i> + <i>nd2</i> )	–5,090.20	18
AIC	10,216.4	
Independent origin		
InL ( <i>cytb</i> )	–2,503.82	12
InL ( <i>nd2</i> )	–2,608.17	12
Total InL	–5,111.99	24
AIC	10,271.97	

Nucleotide sequences of the mitochondrial genes *cytb* and *nd2* from cow, deer and hippopotamus were analysed by PAML<sup>10</sup> with the GTR +  $\Gamma$  model assuming the relations of ((cow, deer), hippopotamus) for the common origin model. The 5'-terminal 1,038 bp (excluding the initiation codon) were used without making further alignment between the two different genes. The common origin model gave a lower AIC value than the independent origin model. InL, log-likelihood score.

1. Darwin, C. *On the Origin of Species by Means of Natural Selection, or, The Preservation of Favoured Races in the Struggle for Life* (J. Murray, 1859).
2. Theobald, D. L. A formal test of the theory of universal common ancestry. *Nature* **465**, 219–222 (2010).
3. Akaike, H. in *Second International Symposium on Information Theory* (eds Petrov, B. N. & Csaki, F.) 267–281 (Akademai Kiado, 1973).
4. Dawkins, R. *The Blind Watchmaker: Why the Evidence of Evolution Reveals a Universe without Design* (Longman Scientific & Technical, 1986).
5. Stewart, C.-B., Schilling, J. W. & Wilson, A. C. Adaptive evolution in the stomach lysozymes of foregut fermenters. *Nature* **330**, 401–404 (1987).
6. Christin, P. A., Salamin, N., Savolainen, V., Duvall, M. R. & Besnard, G. C<sub>4</sub> photosynthesis evolved in grasses via parallel adaptive genetic changes. *Curr. Biol.* **17**, 1241–1247 (2007).
7. Liu, Y. *et al.* Convergent sequence evolution between echolocating bats and dolphins. *Curr. Biol.* **20**, R53–R54 (2010).
8. Li, Y., Liu, Z., Shi, P. & Zhang, J. The hearing gene Prestin unites echolocating bats and whales. *Curr. Biol.* **20**, R55–R56 (2010).
9. Zhong, B., Yonezawa, T., Zhong, Y. & Hasegawa, M. The position of Gnetales among seed plants: overcoming pitfalls of chloroplast phylogenomics. *Mol. Biol. Evol.* doi:10.1093/molbev/msq170 (2 July 2010).
10. Yang, Z. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).

**Author Contributions** T.Y. and M.H. designed the work, T.Y. carried out the analysis, and M.H. wrote the manuscript.

**Competing financial interests:** declared none.

doi:10.1038/nature09482

## Theobald reply

REPLYING TO T. Yonezawa & M. Hasegawa *Nature* **468**, doi:10.1038/nature09482 (2010)

Yonezawa and Hasegawa<sup>1</sup> provide an example from two apparently unrelated families of nucleic acid coding sequences for which an Akaike information criterion (AIC) model selection test, similar to mine<sup>2</sup>, chooses a common origin hypothesis. Although this may seem surprising, the coding sequences in this example were aligned in the same reading frame. The constraints of the genetic code are expected to induce correlations between these sequences (and among all coding sequences) that are not due to common ancestry. For instance, owing to codon bias and the structure of the genetic code, in these sequences the second codon position is biased towards T (about twofold over average), whereas the third position is usually an A (~50%) and rarely a G (~4%).

One can account for these correlations explicitly by using codon models (as implemented in PAML<sup>3</sup>, codonFreq = 2 or 3) or standard amino acid models (as in PhyML<sup>4</sup>). With these more realistic models, independent ancestry is the strongly preferred hypothesis. Furthermore, the raw likelihoods and AIC scores increase significantly (by hundreds to thousands of logs), indicating that codon and amino acid models are greatly superior to the naive nucleotide models.

Yonezawa and Hasegawa<sup>1</sup> point out that I<sup>2</sup> did not explicitly test models in which selection or biophysical constraints generate sequence correlations among proteins with independent origins. Formal phylogenetic models accounting for such factors are currently unavailable; their development would be a welcome advance. Although these are important considerations for proteins with low sequence similarity, neither selection nor physical constraints alone can plausibly generate the high levels of sequence similarity (>55% average sequence identity) observed in the universal protein data set that I used<sup>2,5</sup>. The amount of adaptive convergence necessary to produce thousands of identical amino acids among 23 different proteins from completely independent beginnings is not comparable to the limited molecular convergence seen with, for example, homologous digestive lysozymes<sup>6</sup>, in which already highly similar proteins (in function, structure and sequence) later acquired a handful of identical substitutions in parallel.

How could selection or biophysical constraints induce correlations among unrelated sequences? If certain similar amino acid sequences are necessary for performing specific functions (or for adopting a specific tertiary conformation that is necessary for function), then selection for function may 'lead' proteins with independent origins to neighbouring regions of sequence space. However, no particular protein sequence or fold is necessary for any given function. There are abundant examples of proteins with undetectable sequence similarity and different folds that perform the same biochemical and cellular functions<sup>7</sup>. For example, the proteases subtilisin, trypsin and carboxypeptidase have the same active site and mechanism, whereas papain, renin and thermolysin have different active sites and different mechanisms. All six proteases have radically different folds and sequences. Because different folds in general have different sequence requirements, proteins with the same function need not have similar sequences.

Even assuming that a certain protein fold is necessary for a given function, current molecular evidence indicates that sequence requirements for a fold are extremely low—nearly indistinguishable from random. This data comes from many independent sources from throughout biology.

Many large classes of proteins with identical folds have no detectable sequence similarity (for example, families of TIM barrels, carbonic

anhydrases, OB-folds, SH3 domains, Rossmann folds and immunoglobulin domains). These proteins provide *prima facie* evidence that sequence requirements for any particular fold and function are nearly indistinguishable from random. Protein domains in the SCOP database<sup>8</sup> from different superfamilies yet with the same fold share ~9% sequence identity<sup>9</sup>.

Identical folds with known independent origins have nearly random sequence similarity<sup>9,10</sup>. For example, unrelated proteins with the same fold from the MALISAM database share  $8.5 \pm 0.4\%$  sequence identity<sup>9,10</sup>. This data can be used to estimate the correlations among independently evolved and created proteins with the same fold, and the correlations are nearly random. In the universal protein data set that I used<sup>2</sup>, the average sequence correlation induced by common ancestry is roughly one log-likelihood per site for the most divergent proteins. In contrast, the correlations among independent proteins with the same fold are ~100 times weaker. From this we can estimate that model selection scores for common ancestry hypotheses will be many thousands of logs greater than competing selection hypotheses.

Even the most conserved proteins have not yet reached the limits of sequence space, which has been estimated to be near the random expectation for any given fold and function<sup>11</sup>.

These arguments are largely circumstantial and informal. I have not tested all possible competing hypotheses, and my analysis will not be the "last word on common ancestry"<sup>12</sup>. I emphasize that I have in no sense provided an absolute 'proof' of universal common ancestry. One of the great advantages of the model selection framework that I presented is that if a novel model is proposed with a well-defined likelihood function, then we can easily compare it to the common ancestry models and see how it fares.

D. L. Theobald<sup>1</sup>

<sup>1</sup>Department of Biochemistry, Brandeis University, Waltham, Massachusetts 01778, USA.  
e-mail: dtheobald@brandeis.edu

1. Yonezawa, T. & Hasegawa, M. Was the universal common ancestry proved? *Nature* **468**, 10.1038/nature09482 (2010).
2. Theobald, D. L. A formal test of the theory of universal common ancestry. *Nature* **465**, 219–222 (2010).
3. Yang, Z. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **13**, 555–556 (1997).
4. Guindon, S. & Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696–704 (2003).
5. Brown, J. R., Douady, C. J., Italia, M. J., Marshall, W. E. & Stanhope, M. J. Universal trees based on large combined protein sequence data sets. *Nature Genet.* **28**, 281–285 (2001).
6. Stewart, C. B., Schilling, J. W. & Wilson, A. C. Adaptive evolution in the stomach lysozymes of foregut fermenters. *Nature* **330**, 401–404 (1987).
7. Omelchenko, M. V., Galperin, M. Y., Wolf, Y. I. & Koonin, E. V. Non-homologous isofunctional enzymes: a systematic analysis of alternative solutions in enzyme evolution. *Biol. Direct* **5**, 31 (2010).
8. Andreeva, A. *et al.* Data growth and its impact on the SCOP database: new developments. *Nucleic Acids Res.* **36**, D419–D425 (2008).
9. Cheng, H., Kim, B. H. & Grishin, N. V. Discrimination between distant homologs and structural analogs: lessons from manually constructed, reliable data sets. *J. Mol. Biol.* **377**, 1265–1278 (2008).
10. Cheng, H., Kim, B. H. & Grishin, N. V. MALISAM: a database of structurally analogous motifs in proteins. *Nucleic Acids Res.* **36**, D211–D217 (2008).
11. Povolotskaya, I. S. & Kondrashov, F. A. Sequence space and the ongoing expansion of the protein universe. *Nature* **465**, 922–926 (2010).
12. Steel, M. & Penny, D. Origins of life: Common ancestry put to the test. *Nature* **465**, 168–169 (2010).

doi:10.1038/nature09483

# H2AX prevents CtIP-mediated DNA end resection and aberrant repair in G1-phase lymphocytes

Beth A. Helmink<sup>1</sup>, Anthony T. Tubbs<sup>1</sup>, Yair Dorsett<sup>1</sup>, Jeffrey J. Bednarski<sup>1,2</sup>, Laura M. Walker<sup>1</sup>, Zhihui Feng<sup>3</sup>, Girdhar G. Sharma<sup>3</sup>, Peter J. McKinnon<sup>4</sup>, Junran Zhang<sup>3</sup>, Craig H. Bassing<sup>5</sup> & Barry P. Sleckman<sup>1</sup>

DNA double-strand breaks (DSBs) are generated by the recombination activating gene (RAG) endonuclease in all developing lymphocytes as they assemble antigen receptor genes<sup>1</sup>. DNA cleavage by RAG occurs only at the G1 phase of the cell cycle and generates two hairpin-sealed DNA (coding) ends that require nucleolytic opening before their repair by classical non-homologous end-joining (NHEJ)<sup>1–3</sup>. Although there are several cellular nucleases that could perform this function, only the Artemis nuclease is able to do so efficiently<sup>2,3</sup>. Here, *in vivo*, we show that in murine cells the histone protein H2AX prevents nucleases other than Artemis from processing hairpin-sealed coding ends; in the absence of H2AX, CtIP can efficiently promote the hairpin opening and resection of DNA ends generated by RAG cleavage. This CtIP-mediated resection is inhibited by  $\gamma$ -H2AX and by MDC-1 (mediator of DNA damage checkpoint 1), which binds to  $\gamma$ -H2AX in chromatin flanking DNA DSBs. Moreover, the ataxia telangiectasia mutated (ATM) kinase activates antagonistic pathways that modulate this resection. CtIP DNA end resection activity is normally limited to cells at post-replicative stages of the cell cycle, in which it is essential for homology-mediated repair<sup>4,5</sup>. In G1-phase lymphocytes, DNA ends that are processed by CtIP are not efficiently joined by classical NHEJ and the joints that do form frequently use micro-homologies and show significant chromosomal deletions. Thus, H2AX preserves the structural integrity of broken DNA ends in G1-phase lymphocytes, thereby preventing these DNA ends from accessing repair pathways that promote genomic instability.

V(D)J recombination, the reaction that assembles the second exon of all antigen receptor genes, requires the generation and repair of DNA DSBs and occurs exclusively during the G1 phase of the cell cycle<sup>1</sup>. The Rag-1 and Rag-2 proteins, which together form an endonuclease referred to as RAG, initiate V(D)J recombination by introducing DNA DSBs at the border of two gene segments and their associated recombination signal (RS) RAG recognition sequences<sup>1</sup>. Cleavage by RAG generates a pair of hairpin-sealed coding ends and a pair of blunt signal ends<sup>1</sup>. These distinct pairs of DNA ends are processed and joined by NHEJ to form a coding joint and a signal joint, respectively<sup>2,3</sup>. Signal ends undergo minimal nucleolytic processing before joining. In contrast, hairpin-sealed coding ends must first be opened by an endonuclease and are also frequently processed by exonucleases before joining<sup>2,3</sup>. This nucleolytic processing results in the antigen receptor gene sequence diversification that is essential for adaptive immunity.

Efficient opening of the hairpin-sealed coding ends generated by RAG cleavage is dependent on the Artemis nuclease<sup>2,3</sup>. Although other cellular nucleases have enzymatic activity that could open and resect hairpin-sealed coding ends, this functional redundancy is not evident in Artemis-deficient cells; as a result, Artemis-deficient mice and humans are severely lymphopenic<sup>3</sup>. Thus, the nucleolytic processing

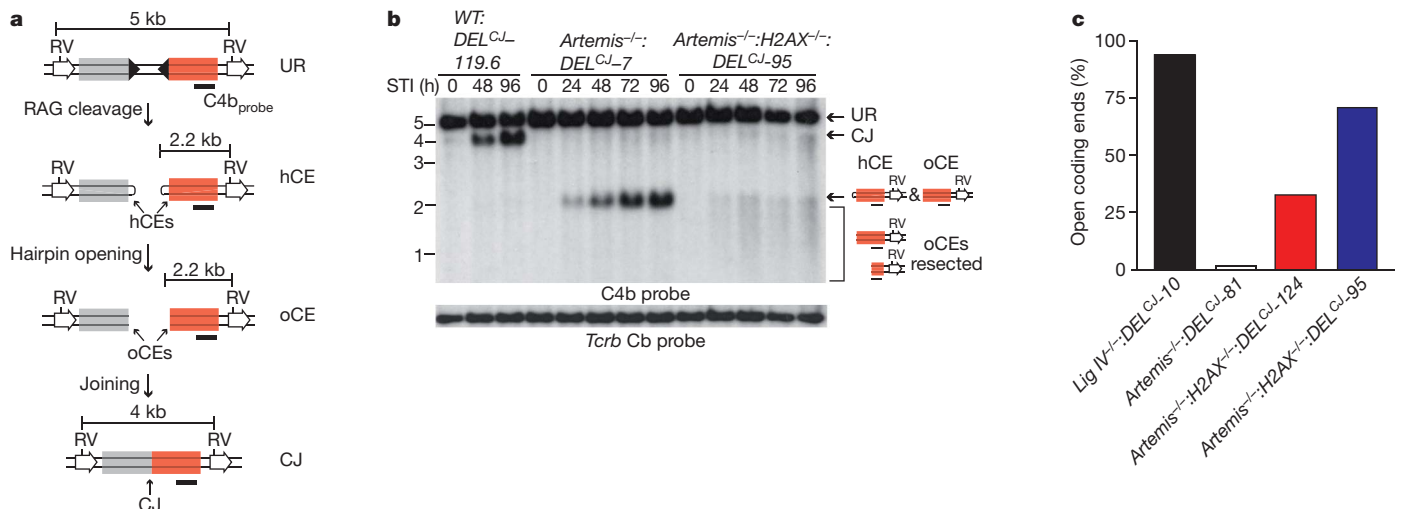
of broken DNA ends in G1-phase lymphocytes must be tightly regulated; however, the cellular components that mediate this regulation are not known. Unrepaired coding ends are resolved as chromosomal translocations at a higher frequency in lymphocytes deficient in both Artemis and H2AX than in lymphocytes with an isolated deficiency of Artemis<sup>6</sup>. Because the formation of these translocations requires the Artemis-independent opening of hairpin-sealed DNA ends, we considered that H2AX might function to restrict the ability of nucleases to act on broken DNA ends in G1-phase lymphocytes.

To test this notion, we generated pre-B-cell lines transformed with the v-abl kinase (hereafter referred to as abl pre-B cells), each of which contained a single integrant of the pMX-DEL<sup>CJ</sup> retroviral recombination substrate and was wild type (WT:DEL<sup>CJ</sup>), deficient in Artemis (Artemis<sup>-/-</sup>:DEL<sup>CJ</sup>) or deficient in both Artemis and H2AX (Artemis<sup>-/-</sup>:H2AX<sup>-/-</sup>:DEL<sup>CJ</sup>)<sup>6–8</sup>. pMX-DEL<sup>CJ</sup> had a single pair of recombination signals and rearranged by deletion, resulting in a coding joint that remained within the chromosomal context (Fig. 1a)<sup>7</sup>. Treatment of abl pre-B cells with the v-abl kinase inhibitor STI571 leads to cell cycle arrest in G1 and the induction of RAG, which in NHEJ-deficient cells results in the accumulation of unrepaired coding ends<sup>6–9</sup>. In G1-arrested Artemis<sup>-/-</sup>:DEL<sup>CJ</sup> abl pre-B cells, these unrepaired coding ends were homogeneous in size, as expected given their hairpin-sealed structure (Fig. 1b and Supplementary Figs 1 and 2). In contrast, coding ends in G1-arrested Artemis<sup>-/-</sup>:H2AX<sup>-/-</sup>:DEL<sup>CJ</sup> abl pre-B cells were heterogeneous in size and significantly smaller (up to 1 kilobase), suggesting that these hairpin-sealed coding ends had been opened and resected in an Artemis-independent fashion (Fig. 1b and Supplementary Figs 1 and 2). H2AX also regulates coding-end resection in DNA ligase IV-deficient (LigIV<sup>-/-</sup>) abl pre-B cells, in which hairpin-sealed coding ends can be efficiently opened by Artemis but cannot be ligated (Supplementary Fig. 3). Moreover, blunt chromosomal signal ends generated by RAG cleavage at the pMX-DEL<sup>SJ</sup> retroviral substrate also show significant resection in LigIV<sup>-/-</sup>:H2AX<sup>-/-</sup>:DEL<sup>SJ</sup> abl pre-B cells (Supplementary Fig. 4)<sup>7</sup>.

To show that hairpin-sealed coding ends have been opened, DNA end structure was assayed by TdT-assisted PCR, which can detect hairpin-opened but not hairpin-sealed coding ends (Supplementary Fig. 5a)<sup>10</sup>. In this regard, analysis of Artemis<sup>-/-</sup>:H2AX<sup>-/-</sup>:DEL<sup>CJ</sup> abl pre-B cells revealed robust TdT-assisted PCR products, indicating that many of the hairpin-sealed coding ends had been opened (Supplementary Fig. 5c). To quantify the fraction of open coding ends in Artemis<sup>-/-</sup>:H2AX<sup>-/-</sup>:DEL<sup>CJ</sup> abl pre-B cells, we performed Southern blot analyses of native and denatured genomic DNA. Denaturing hairpin-opened coding ends dissociates the complementary DNA strands into single-stranded fragments that migrate differently from hairpin-sealed coding ends, whose complementary strands are covalently linked and therefore do not dissociate (Supplementary Fig. 6). Analyses of denatured coding ends from Artemis<sup>-/-</sup>:H2AX<sup>-/-</sup>:DEL<sup>CJ</sup>

<sup>1</sup>Department of Pathology and Immunology, Washington University School of Medicine, St Louis, Missouri 63110, USA. <sup>2</sup>Department of Pediatrics, Washington University School of Medicine, St Louis, Missouri 63110, USA. <sup>3</sup>Department of Radiation Oncology, Washington University School of Medicine, St Louis, Missouri 63110, USA. <sup>4</sup>Department of Genetics and Tumor Cell Biology, St Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. <sup>5</sup>Division of Cancer Pathobiology, Department of Pathology and Laboratory Medicine, Center for Childhood Cancer Research, Children's Hospital of Philadelphia, University of Pennsylvania School of Medicine, Abramson Family Cancer Research Institute, Philadelphia, Pennsylvania 19104, USA.





**Figure 1 | H2AX inhibits DNA end resection in G1-phase lymphocytes.**

**a**, pMX-DEL<sup>CJ</sup> retroviral recombination substrate unrearranged (UR), with coding ends that are hairpin-sealed (hCE) or open (oCE), and coding joint (CJ). Recombination signals (filled triangles), EcoRV sites (RV), the C4b probe (black bar) and fragment sizes are shown. **b**, Southern blot analysis of EcoRV-digested genomic DNA from wild-type (WT:DEL<sup>CJ</sup>-119.6), Artemis<sup>-/-</sup>:DEL<sup>CJ</sup> (Artemis<sup>-/-</sup>:DEL<sup>CJ</sup>-7) and Artemis<sup>-/-</sup>:H2AX<sup>-/-</sup>:DEL<sup>CJ</sup> (Artemis<sup>-/-</sup>:H2AX<sup>-/-</sup>:DEL<sup>CJ</sup>-95) abl pre-B cell lines treated with STI571

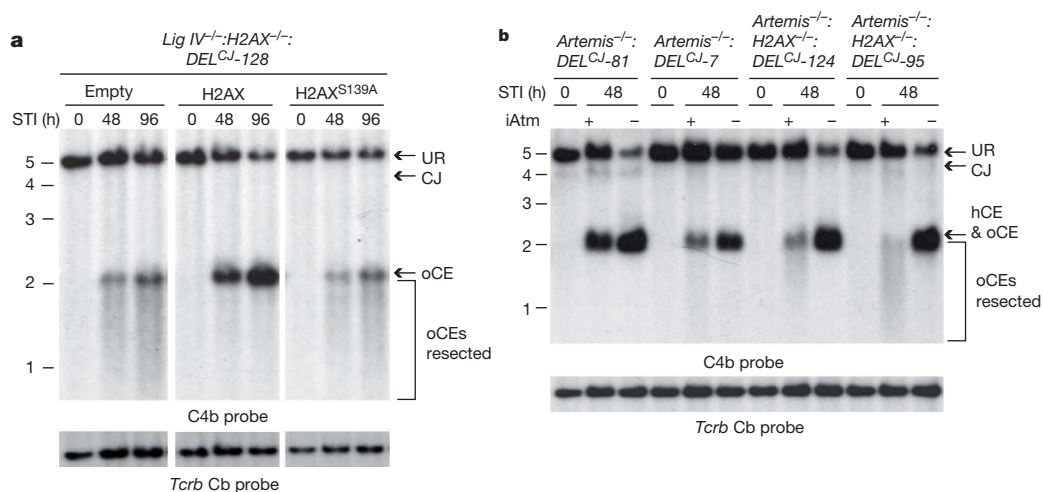
(STI). C4b hybridizing UR, CJ and hCEs are indicated, as are oCEs that have been resected (bracket). Tcrb constant region (Cb) probe hybridization was used as a loading control. **c**, Quantification of denaturing Southern blot analyses for open coding ends in LigIV<sup>-/-</sup>:DEL<sup>CJ</sup> (LigIV<sup>-/-</sup>:DEL<sup>CJ</sup>-10), Artemis<sup>-/-</sup>:DEL<sup>CJ</sup> (Artemis<sup>-/-</sup>:DEL<sup>CJ</sup>-81) and Artemis<sup>-/-</sup>:H2AX<sup>-/-</sup>:DEL<sup>CJ</sup> (Artemis<sup>-/-</sup>:H2AX<sup>-/-</sup>:DEL<sup>CJ</sup>-95 and Artemis<sup>-/-</sup>:H2AX<sup>-/-</sup>:DEL<sup>CJ</sup>-124) abl pre-B cells (see Supplementary Figs 2 and 3 for primary data).

abl pre-B cells revealed that up to 70% of the hairpin-sealed coding ends had been opened (Fig. 1c and Supplementary Fig. 6).

The requirement for H2AX in regulating coding-end processing was also observed at the immunoglobulin light chain (*Igl*)*k* locus in primary G1-phase bone-marrow-derived pre-B cell cultures from Artemis<sup>-/-</sup> and Artemis<sup>-/-</sup>:H2AX<sup>-/-</sup> mice expressing an immunoglobulin heavy chain (*Igh*) transgene (*Ightg*) (Supplementary Fig. 7)<sup>8</sup>. Southern blotting and TdT-assisted PCR revealed that the Jk coding ends in Artemis<sup>-/-</sup>:H2AX<sup>-/-</sup>:*Ightg* primary pre-B cells had open hairpins and were resected, whereas those from Artemis<sup>-/-</sup>:*Ightg* pre-B cells remained hairpin-sealed (Supplementary Fig. 8). Taken together, these data show that H2AX restricts the activity of nucleolytic pathways that would otherwise aberrantly resect unrepaired hairpin-sealed or open coding ends in lymphocytes at the G1-phase of the cell cycle.

H2AX-dependent DNA damage responses generally depend on the phosphorylation of serine 139 by ATM or DNAPKcs to form

γ-H2AX in chromatin flanking DNA DSBs, including those generated by RAG<sup>11,12</sup>. Reconstitution of LigIV<sup>-/-</sup>:H2AX<sup>-/-</sup>:DEL<sup>CJ</sup> and Artemis<sup>-/-</sup>:H2AX<sup>-/-</sup>:DEL<sup>CJ</sup> abl pre-B cells with wild-type H2AX, but not with a serine 139-to-alanine mutant (H2AX<sup>S139A</sup>), inhibited coding-end resection in these cells (Fig. 2a and Supplementary Fig. 9). These findings implicate γ-H2AX formation at broken DNA ends in maintaining the structure of these ends in G1-phase lymphocytes but do not exclude the possibility that H2AX also inhibits resection through additional pathways that are independent of γ-H2AX formation. Although ATM is required for the optimal formation of γ-H2AX in chromatin flanking RAG DSBs, treatment of Artemis<sup>-/-</sup>:DEL<sup>CJ</sup> abl pre-B cells with the ATM kinase inhibitor KU-55933 did not lead to the opening and resection of hairpin-sealed coding ends (Fig. 2b)<sup>12</sup>. Rather, treatment of Artemis<sup>-/-</sup>:H2AX<sup>-/-</sup>:DEL<sup>CJ</sup> abl pre-B cells with KU-55933 resulted in a significant block in end resection, with a large fraction of DNA ends in these cells being hairpin-sealed (Fig. 2b and



**Figure 2 | ATM and γ-H2AX regulate DNA end resection.** **a**, pMX-DEL<sup>CJ</sup> Southern blot analysis (as described in Fig. 1a, b) of STI571-treated LigIV<sup>-/-</sup>:H2AX<sup>-/-</sup>:DEL<sup>CJ</sup> abl pre-B cells reconstituted with an empty retroviral vector (empty) or vectors encoding wild-type H2AX (H2AX) or

H2AX<sup>S139A</sup>. **b**, pMX-DEL<sup>CJ</sup> Southern blot analysis of Artemis<sup>-/-</sup>:DEL<sup>CJ</sup> and Artemis<sup>-/-</sup>:H2AX<sup>-/-</sup>:DEL<sup>CJ</sup> abl pre-B-cell lines treated with either the ATM inhibitor (iATM) KU55933 (+) or a dimethylsulphoxide vehicle control (-).

data not shown). We conclude that ATM functions to inhibit coding-end resection through the formation of  $\gamma$ -H2AX; however, ATM activity is also required to promote this nucleolytic resection. Thus, ATM modulates the function of antagonistic pathways that both positively and negatively regulate DNA end resection in G1-phase lymphocytes.

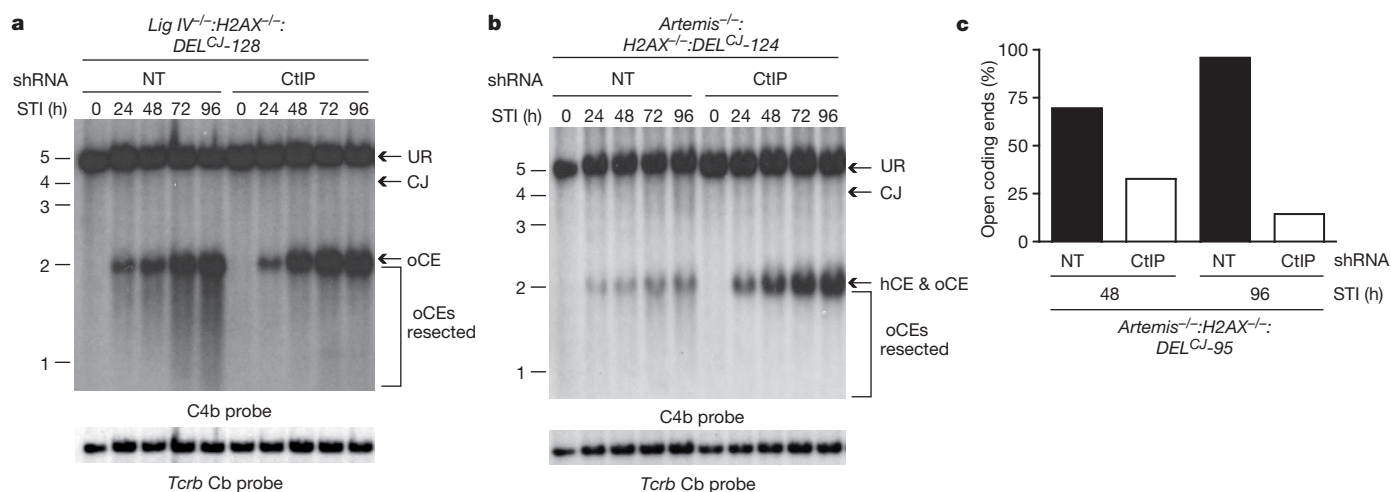
We show that, in the absence of H2AX, nucleases other than Artemis can efficiently open and resect hairpin-sealed coding ends in a manner that is dependent on ATM. In this regard, ATM positively regulates CtIP activity in promoting DNA end resection<sup>5</sup>. CtIP binds directly to Nbs1, a component of the Mre11–Rad50–Nbs1 (MRN) complex that associates with RAG DSBs and functions in their repair<sup>5,11,13–15</sup>. Furthermore, Sae2, the *S. cerevisiae* orthologue of CtIP, functions with Mre11 to promote the opening and resection of hairpin-sealed DNA ends *in vitro*<sup>16</sup>. Taken together, these data suggest that H2AX could regulate the ability of CtIP to mediate hairpin opening and DNA end resection in G1-phase lymphocytes. Indeed, knockdown of CtIP in both *Artemis*<sup>-/-</sup>:*H2AX*<sup>-/-</sup>:*DEL*<sup>CJ</sup> and *Lig IV*<sup>-/-</sup>:*H2AX*<sup>-/-</sup>:*DEL*<sup>CJ</sup> abl pre-B cells largely blocked the aberrant coding-end resection observed in these cells (Fig. 3a, b and Supplementary Figs 10 and 11). Moreover, in CtIP-deficient *Artemis*<sup>-/-</sup>:*H2AX*<sup>-/-</sup>:*DEL*<sup>CJ</sup> abl pre-B cells most unrepaired coding ends were hairpin-sealed (Fig. 3c). Although our data show that the hairpin coding ends in *Artemis*<sup>-/-</sup>:*H2AX*<sup>-/-</sup>:*DEL*<sup>CJ</sup> abl pre-B cells have been opened, we cannot determine the position of opening or the extent of resection after opening. However, it is notable that Sae2 can mediate hairpin opening at significant distances from the hairpin tip *in vitro*<sup>16</sup>.

The DNA-damage-response protein MDC-1 is recruited by  $\gamma$ -H2AX to chromatin flanking DNA DSBs<sup>17,18</sup>. We found that, like  $\gamma$ -H2AX, MDC-1 was also required for the inhibition of the ATM-dependent resection of coding ends in G1-phase lymphocytes (Supplementary Fig. 12). Because CtIP binds to the FHA domain of Nbs1 that also binds MDC-1, this raises the possibility that, on recruitment to DSBs by  $\gamma$ -H2AX, MDC-1 may disrupt CtIP–Nbs1 interactions<sup>19–23</sup>. The DNA-damage-response protein 53BP1, which is also retained at DSB sites in a  $\gamma$ -H2AX-dependent manner, regulates DNA end resection during V(D)J recombination in thymocytes and during immunoglobulin class switch recombination<sup>10,24</sup>. In addition, 53BP1 inhibits CtIP-dependent DNA end resection in Brca-1-deficient cells at post-replicative stages of the cell cycle; thus,  $\gamma$ -H2AX may inhibit DNA end resection in these cells by recruiting 53BP1 (ref. 25). However, at post-replicative stages of the cell cycle, H2AX also functions to promote DNA DSB repair by homologous recombination, which requires CtIP-mediated DNA end resection<sup>5,26</sup>. Thus, H2AX

may function in a cell-cycle-specific manner to modulate the activity of several pathways that regulate DNA end resection.

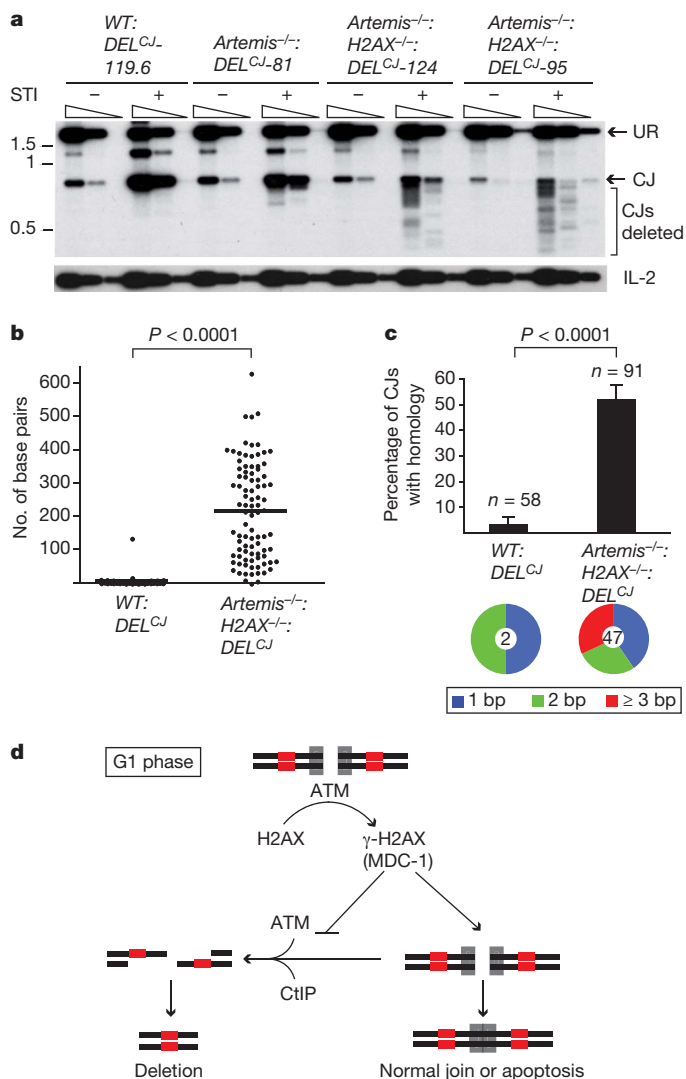
What is the fate of broken DNA ends processed by CtIP in G1-phase lymphocytes? During V(D)J recombination, Artemis functions primarily to open hairpin-sealed coding ends, after which core NHEJ factors join these DNA ends<sup>3</sup>. However, hairpin-sealed coding ends opened in a CtIP-dependent manner persisted unrepaired at high levels in *Artemis*<sup>-/-</sup>:*H2AX*<sup>-/-</sup> abl pre-B cells (Fig. 1b and Supplementary Fig. 2). In this regard, the single-strand overhangs generated by CtIP-mediated resection during homologous recombination would probably be poor substrates for NHEJ in G1-phase cells<sup>4</sup>. However, this resection could expose regions of homology flanking the DSB, which, if used to mediate DSB repair by homology-driven repair pathways, would form joints with chromosomal deletions<sup>4</sup>. Indeed, PCR and sequence analyses of coding joints formed in *Artemis*<sup>-/-</sup>:*H2AX*<sup>-/-</sup>:*DEL*<sup>CJ</sup> abl pre-B cells revealed that they were heterogeneous in size, in contrast with those formed in WT:*DEL*<sup>CJ</sup> or *Artemis*<sup>-/-</sup>:*DEL*<sup>CJ</sup> abl pre-B cells (Fig. 4a, b and Supplementary Figs 13 and 14). These joints had significant deletions extending up to 700 base pairs, which is the maximum size deletion that would be amplified by the PCR approach used (Fig. 4a, b and Supplementary Fig. 14). Moreover, the coding joints formed in *Artemis*<sup>-/-</sup>:*H2AX*<sup>-/-</sup>:*DEL*<sup>CJ</sup> abl pre-B cells used microhomologies at a higher frequency than those formed in WT:*DEL*<sup>CJ</sup> abl pre-B cells (50% versus 5%; Fig. 4c and Supplementary Fig. 14). Analysis of T-cell receptor  $\beta$  (*Tcrb*) Db1 to Jb1.1/Jb1.2 joints in *Artemis*<sup>-/-</sup>:*H2AX*<sup>-/-</sup> thymocytes revealed that these joints similarly showed significant deletions compared with those formed in either *Artemis*<sup>-/-</sup> or wild-type thymocytes (Supplementary Fig. 15). We conclude that RAG-mediated DNA breaks generated in H2AX-deficient lymphocytes that are processed in a CtIP-dependent manner are resistant to repair by classical NHEJ. However, these DNA ends can be channelled into homology-driven repair pathways, resulting in joints that form significant chromosomal deletions.

The requirement for H2AX in the prevention of CtIP-dependent resection and the resolution of RAG DSBs as chromosomal deletions is congruent with the phenotype of H2AX-deficient mice, which are predisposed to lymphoid tumours that can harbour chromosomal lesions indicative of aberrantly resolved RAG DSBs<sup>27,28</sup>. However, chromosomal V(D)J recombination proceeds efficiently in H2AX-deficient abl pre-B cells<sup>6,29</sup>. Moreover, DbJb joints formed in *H2AX*<sup>-/-</sup> thymocytes did not exhibit significant deletions or an increase in micro-homology usage in comparison with wild-type thymocytes (Supplementary Figs 16 and 17). Thus, H2AX may be required for



**Figure 3 | H2AX prevents CtIP-mediated DNA end resection.** **a, b**, pMX-*DEL*<sup>CJ</sup> Southern blot analysis (as described in Fig. 1a, b) of STI571-treated *Lig IV*<sup>-/-</sup>:*H2AX*<sup>-/-</sup>:*DEL*<sup>CJ</sup> (**a**) and *Artemis*<sup>-/-</sup>:*H2AX*<sup>-/-</sup>:*DEL*<sup>CJ</sup> (**b**) abl pre-B cells expressing either non-targeting (NT) or CtIP-specific (CtIP) shRNAs.

**(c)** Quantification of denaturing Southern blot analysis for open CEs from STI571-treated *Artemis*<sup>-/-</sup>:*H2AX*<sup>-/-</sup>:*DEL*<sup>CJ</sup> abl pre-B cells expressing either NT or CtIP-specific shRNAs.



**Figure 4 | Aberrant joining in H2AX-deficient cells.** **a**, PCR products representing normal and deleted (bracketed) pMX- $DEL^{CJ}$  CJs from WT: $DEL^{CJ}$ ,  $Artemis^{-/-}$ : $DEL^{CJ}$  and  $Artemis^{-/-}$ : $H2AX^{-/-}$ : $DEL^{CJ}$  abl pre-B cells treated with STI571. Serial fivefold dilutions of genomic DNA were amplified. IL-2 gene PCR was used as a loading control. **b**, **c**, Base-pairs deleted (**b**) and microhomology utilization (**c**) in pMX- $DEL^{CJ}$  CJs sequenced from WT: $DEL^{CJ}$  and  $Artemis^{-/-}$ : $H2AX^{-/-}$ : $DEL^{CJ}$  abl pre-B cells (Supplementary Fig. 14). The total number of sequenced coding joints ( $n$ ) is indicated in **c**. The pie chart shows the fraction of joints with 1, 2 or at least 3 microhomologies and the total number (centre) of joints with microhomologies. **d**, Model for H2AX function in DNA end processing in G1-phase lymphocytes. Red blocks represent homologous sequences. Results are shown as means and s.e.m.;  $P$  values were determined by using Student's  $t$ -test with Welch's correction for unequal variances.

the repair of a limited subset of RAG DSBs. Alternatively, other proteins may compensate for a more general requirement for H2AX in DNA DSB repair during V(D)J recombination. In agreement with this notion, whereas isolated deficiencies in H2AX or XRCC4-like factor (XLF) have a minimal effect on chromosomal V(D)J recombination, the combined deficiency of H2AX and XLF in abl pre-B cells results in a significant accumulation of unrepaired coding ends<sup>29</sup>. Moreover, these coding ends are more extensively resected than those in either  $Artemis^{-/-}$ : $H2AX^{-/-}$  or  $Lig IV^{-/-}$ : $H2AX^{-/-}$  abl pre-B cells, raising the possibility that H2AX and XLF have overlapping activities in modulating DNA end resection<sup>29</sup>.

We have shown that H2AX maintains the structural integrity of broken DNA ends generated during V(D)J recombination in G1-phase lymphocytes. This protective function of H2AX would ensure

that unrepaired DNA ends are either joined through classical NHEJ or signal for elimination of the cell by apoptosis (Fig. 4d). In the absence of H2AX, ATM/CtIP-dependent resection creates genomic instability by allowing DNA ends to be shuttled into homology-driven repair pathways that can form potentially dangerous chromosomal deletions (Fig. 4d). Formation of chromosomal translocations in H2AX-deficient mice may rely on these defects in DNA end processing coupled with the diminished retention of coding ends in post-cleavage complexes in H2AX-deficient cells<sup>6</sup>. Significant parallels can be drawn between mechanisms that protect, process and repair RAG DSBs and those incurred by environmental genotoxins. In this regard, a subset of genotoxic DSBs require Artemis for their repair by NHEJ, suggesting that these broken DNA ends must undergo nucleolytic processing<sup>30</sup>. Furthermore, the repair of these genotoxic DSBs may also depend on H2AX<sup>30</sup>. Thus, our finding that H2AX regulates the processing of unrepaired DNA ends during V(D)J recombination may reflect a broader function of H2AX in regulating the nucleolytic processing of DNA DSBs generated during other physiological processes and by genotoxic agents.

## METHODS SUMMARY

Abl pre-B cell lines containing the pMX- $DEL^{CJ}$  retroviral recombination substrate were generated and maintained, and interleukin (IL)-7-dependent pre-B cells were cultured as described previously<sup>7,8</sup>. Standard protocols for Southern blotting (native and denaturing), western blotting, flow cytometry, retroviral-mediated protein expression and short hairpin RNA (shRNA)-mediated knockdown were followed. TdT-assisted PCR for coding ends generated during rearrangement of pMX- $DEL^{CJ}$  and the *Igk* locus was performed as described in Methods<sup>10</sup>.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 14 June; accepted 15 October 2010.**

**Published online 15 December 2010.**

1. Fugmann, S. D., Lee, A. I., Shockett, P. E., Villy, I. J. & Schatz, D. G. The RAG proteins and V(D)J recombination: complexes, ends, and transposition. *Annu. Rev. Immunol.* **18**, 495–527 (2000).
2. Lieber, M. R., Ma, Y., Pannicke, U. & Schwarz, K. The mechanism of vertebrate nonhomologous DNA end joining and its role in V(D)J recombination. *DNA Repair (Amst.)* **3**, 817–826 (2004).
3. Rooney, S., Chaudhuri, J. & Alt, F. W. The role of the non-homologous end-joining pathway in lymphocyte development. *Immunol. Rev.* **200**, 115–131 (2004).
4. Huertas, P. DNA resection in eukaryotes: deciding how to fix the break. *Nature Struct. Mol. Biol.* **17**, 11–16 (2010).
5. You, Z. & Bailis, J. M. DNA damage and decisions: CtIP coordinates DNA repair and cell cycle checkpoints. *Trends Cell Biol.* **20**, 402–409 (2010).
6. Yin, B. *et al.* Histone H2AX stabilizes broken DNA strands to suppress chromosome breaks and translocations during V(D)J recombination. *J. Exp. Med.* **206**, 2625–2639 (2009).
7. Bredemeyer, A. L. *et al.* ATM stabilizes DNA double-strand-break complexes during V(D)J recombination. *Nature* **442**, 466–470 (2006).
8. Bredemeyer, A. L. *et al.* DNA double-strand breaks activate a multi-functional genetic program in developing lymphocytes. *Nature* **456**, 819–823 (2008).
9. Muljo, S. A. & Schlissel, M. S. A small molecule Abl kinase inhibitor induces differentiation of Abelson virus-transformed pre-B cell lines. *Nature Immunol.* **4**, 31–37 (2003).
10. Difilippantonio, S. *et al.* 53BP1 facilitates long-range DNA end-joining during V(D)J recombination. *Nature* **456**, 529–533 (2008).
11. Chen, H. T. *et al.* Response to RAG-mediated VDJ cleavage by NBS1 and  $\gamma$ -H2AX. *Science* **290**, 1962–1965 (2000).
12. Savic, V. *et al.* Formation of dynamic  $\gamma$ -H2AX domains along broken DNA strands is distinctly regulated by ATM and MDC1 and dependent upon H2AX densities in chromatin. *Mol. Cell* **34**, 298–310 (2009).
13. Chen, L., Nievera, C. J., Lee, A. Y. & Wu, X. Cell cycle-dependent complex formation of BRCA1/CtIP/MRN is important for DNA double-strand break repair. *J. Biol. Chem.* **283**, 7713–7720 (2008).
14. Deriano, L., Stracker, T. H., Baker, A., Petrini, J. H. & Roth, D. B. Roles for NBS1 in alternative nonhomologous end-joining of V(D)J recombination intermediates. *Mol. Cell* **34**, 13–25 (2009).
15. Helmink, B. A. *et al.* MRN complex function in the repair of chromosomal RAG-mediated DNA double-strand breaks. *J. Exp. Med.* **206**, 669–679 (2009).
16. Lengsfeld, B. M., Rattray, A. J., Bhaskara, V., Ghirlando, R. & Paull, T. T. Sae2 is an endonuclease that processes hairpin DNA cooperatively with the Mre11/Rad50/Xrs2 complex. *Mol. Cell* **28**, 638–651 (2007).
17. Stucki, M. *et al.* MDC1 directly binds phosphorylated histone H2AX to regulate cellular responses to DNA double-strand breaks. *Cell* **123**, 1213–1226 (2005).



18. Lou, Z. *et al.* MDC1 maintains genomic stability by participating in the amplification of ATM-dependent DNA damage signals. *Mol. Cell* **21**, 187–200 (2006).
19. Melander, F. *et al.* Phosphorylation of SPT repeats in the MDC1 N terminus triggers retention of NBS1 at the DNA damage-modified chromatin. *J. Cell Biol.* **181**, 213–226 (2008).
20. Chapman, J. R. & Jackson, S. P. Phospho-dependent interactions between NBS1 and MDC1 mediate chromatin retention of the MRN complex at sites of DNA damage. *EMBO Rep.* **9**, 795–801 (2008).
21. Spycher, C. *et al.* Constitutive phosphorylation of MDC1 physically links the MRE11–RAD50–NBS1 complex to damaged chromatin. *J. Cell Biol.* **181**, 227–240 (2008).
22. Lloyd, J. *et al.* A supramodular FHA/BRCT-repeat architecture mediates Nbs1 adaptor function in response to DNA damage. *Cell* **139**, 100–111 (2009).
23. Williams, R. S. *et al.* Nbs1 flexibly tethers Ctp1 and Mre11–Rad50 to coordinate DNA double-strand break processing and repair. *Cell* **139**, 87–99 (2009).
24. Bothmer, A. *et al.* 53BP1 regulates DNA resection and the choice between classical and alternative end joining during class switch recombination. *J. Exp. Med.* **207**, 855–865 (2010).
25. Bunting, S. F. *et al.* 53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks. *Cell* **141**, 243–254 (2010).
26. Xie, A. *et al.* Distinct roles of chromatin-associated proteins MDC1 and 53BP1 in mammalian double-strand break repair. *Mol. Cell* **28**, 1045–1057 (2007).
27. Bassing, C. H. *et al.* Histone H2AX: a dosage-dependent suppressor of oncogenic translocations and tumors. *Cell* **114**, 359–370 (2003).
28. Celeste, A. *et al.* H2AX haploinsufficiency modifies genomic stability and tumor susceptibility. *Cell* **114**, 371–383 (2003).
29. Zha, S. *et al.* XLF has redundant functions with ATM and H2AX in V(D)J recombination and non-homologous DNA end-joining. *Nature* doi:10.1038/nature09604 (this issue).
30. Riballo, E. *et al.* A pathway of double-strand break rejoining dependent upon ATM, Artemis, and proteins locating to  $\gamma$ -H2AX foci. *Mol. Cell* **16**, 715–724 (2004).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank E. Oltz and F. Alt for critical review of the manuscript; R. Baer for providing us with the anti-CtIP antibodies; and G. Longmore and Y. Feng for pFLRU, pHR'  $\Delta$ 8.2R and pCMV-VSVg. This work is supported by National Institutes of Health (NIH) grants AI074953 (B.P.S.), AI47829 (B.P.S.), CA136470 (B.P.S. and C.H.B.), CA125195 (C.H.B.), CA21765 (P.J.McK.) and NS37956 (P.J.McK.). J.J.B. is supported by a NIH Ruth L. Kirschstein National Research Service Award (T32 HD007499) and a Children's Discovery Institute Fellows Award. C.H.B. was a Pew Scholar in the Biomedical Sciences program and is a Leukemia and Lymphoma Society Scholar.

**Author Contributions** B.P.S. and B.A.H. conceived the study and wrote the paper. B.P.S. and C.H.B. designed experiments and interpreted data. B.A.H., A.T.T., Y.D., G.S., J.Z. and J.J.B. designed and performed experiments and interpreted data. Z.F. and L.M.W. performed experiments. P.J.McK. provided important reagents.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to B.P.S. ([sleckman@wustl.edu](mailto:sleckman@wustl.edu)).

## METHODS

**Mice.** All animals were housed in a specific pathogen-free facility at Washington University School of Medicine, and all animal protocols were approved by the Washington University Institutional Animal Care and Use Committee.

**Cell line generation and cell culture.** *Rag*<sup>-/-</sup>, *Artemis*<sup>-/-</sup> and WT abl pre-B cells were described previously<sup>7,8</sup>. *Artemis*<sup>-/-</sup>:*H2AX*<sup>-/-</sup> v-abl-transformed pre-B cells were generated as described previously<sup>6</sup>. *LigIV*<sup>-/-</sup> and *LigIV*<sup>-/-</sup>:*H2AX*<sup>-/-</sup> abl pre-B cells were generated by treating *LigIV*<sup>loxp/loxp</sup> and *LigIV*<sup>loxp/loxp</sup>:*H2AX*<sup>-/-</sup> abl pre-B cells with a Tat-Cre fusion protein<sup>6,31</sup>. Cells were incubated for 1 h in medium containing 50 µg ml<sup>-1</sup> Tat-Cre and subcloned 48 h later. Cre-mediated deletion of the *DNA Ligase IV* gene was confirmed by PCR and Southern blotting. All lines were transduced with pMX-DEL<sup>CJ</sup> or pMX-DEL<sup>SJ</sup> by co-centrifugation, and clonal populations each containing a single integrant of the retroviral substrate were isolated by limiting dilution as described previously<sup>8</sup>. Cells were treated with 3 µM STI571 (Novartis) for the indicated durations at a concentration of 10<sup>6</sup> cells ml<sup>-1</sup>. KU-55933 (Tocris) was used at a concentration of 15 µM. Primary bone-marrow pre-B-cell cultures were generated by harvesting bone marrow from *Rag*<sup>-/-</sup>:*Ightg*, *Artemis*<sup>-/-</sup>:*Ightg* or *Artemis*<sup>-/-</sup>:*H2AX*<sup>-/-</sup>:*Ightg* mice followed by culture for 6–10 days in medium containing 5 ng ml<sup>-1</sup> IL-7 maintained at a concentration of about 2 × 10<sup>6</sup> cells ml<sup>-1</sup> before withdrawal of IL-7 (ref. 8).

**Retroviral reconstitution and lentiviral knockdown.** Reconstitution of *Artemis*<sup>-/-</sup>:*H2AX*<sup>-/-</sup> and *LigIV*<sup>-/-</sup>:*H2AX*<sup>-/-</sup> abl pre-B cells was performed by retroviral transduction with either empty retrovirus or retrovirus containing cDNAs encoding H2AX or H2AX<sup>S139A</sup>. A cDNA encoding H2AX<sup>S139A</sup> (serine 139 changed to alanine) was generated by PCR-based site-directed mutagenesis of WT H2AX cDNA. cDNAs encoding WT H2AX and H2AX<sup>S139A</sup> were cloned into the pMX-PIE retroviral vector and cells were transduced by co-centrifugation as described previously<sup>8</sup>. Cells expressing the retroviral construct were obtained by flow cytometric cell sorting of cells expressing GFP using a FACSVantage (BD Biosciences). Generation of lentiviral shRNAs vectors was performed with the previously described pFLRU:YFP lentiviral vector<sup>32,33</sup>. CtlP-specific and non-targeting (NT) shRNAs were cloned into the pFLRU:YFP lentiviral vector. Sequences targeted by the shRNA were CtlP (5'-GAGCAGACCTTTCTCAGTA-3') and NT (5'-GGTTCGATGTCCCAATTCTG-3'). These pFLRU:shRNA:YFP vectors (4 µg) were individually co-transfected with 4 µg of pHR<sup>+</sup>Δ8.2R packaging vector and 1 µg of pCMV-VSVg envelope plasmid into HEK293T cells plated at 70% confluence in 6-cm<sup>2</sup> plates, using Lipofectamine 2000 (Invitrogen). Medium was replaced 12 h after transfection. Supernatants were harvested 24 h later. Transduction of abl pre-B cells was performed by co-centrifugation with viral supernatant at 1,800 r.p.m. (650g) for 90 min, with Polybrene added to 5 µg ml<sup>-1</sup>. Cells expressing the pFLRU-shRNA vectors were obtained by flow cytometric cell sorting of cells expressing YFP, with a FACSVantage (BD Biosciences).

**Southern blotting analyses.** Standard Southern blot analysis of pMX-DEL<sup>CJ</sup> and pMX-DEL<sup>SJ</sup> was performed with *EcoRV*-digested genomic DNA and the C4b probe as described previously<sup>8</sup>. Southern blot analyses of coding ends generated during rearrangement at the *Igk* locus were performed as described previously on genomic DNA digested with *SacI* and *EcoRI*, with the JkIII probe<sup>7</sup>. Denaturing agarose-gel electrophoresis was performed as described previously, with modifications<sup>34</sup>. In brief, 40 µg of genomic DNA was digested overnight with *EcoRV* in a 400-µl volume and concentrated to 30 µl. The DNA was then resuspended with the addition of 5 volumes of a solution containing 8 M urea, 1% Nonidet P40, 1 mM Tris-HCl pH 8.0 and 0.5 mg ml<sup>-1</sup> bromophenol blue. This DNA solution was divided in two; one half was heated at 90 °C for 8 min to denature the genomic DNA, and the other half was incubated on ice. After 8 min, the heated DNA samples were placed on ice before electrophoresis at 50 V and 4 °C on a 1.2% agarose Tris-acetate-EDTA gel with 1 M urea for about 24 h in TAE buffer also

containing 1 M urea. Quantification was performed with ImageJ software. To calculate the percentage of hairpin-sealed coding ends, we measured the integrated density (ID) of the band representing open coding ends (oCEs) and that representing hairpin-sealed coding ends (hCEs) and subtracted background ID levels to obtain the corrected ID levels for each (corr. ID<sup>oCE</sup> and corr. ID<sup>hCE</sup>, respectively). The ID for the closed ends was divided by the sum of the IDs of closed and open ends and converted to a percentage: hCE (%) = 100(corr. ID<sup>hCE</sup>)/(corr. ID<sup>hCE</sup> + corr. ID<sup>oCE</sup>). **TdT-assisted PCR.** TdT-assisted PCR analysis of pMX-DEL<sup>CJ</sup> coding ends was performed as described previously, with the IRES REV5 oligonucleotide (5'-CTCGACTAAACACATGTAAAGC-3') for the primary PCR reaction, the IRES REV4 oligonucleotide (5'-CCCTTGTGTGAATACGCTTG-3') for the secondary PCR reaction and the I4 oligonucleotide (5'-TAAGATACACCTGCAAAGGCG-3') as a probe<sup>10</sup>. Similar conditions were used for PCR analyses of coding ends generated during rearrangement of the endogenous *Igk* locus, using the Jk2 ds oligonucleotide (5'-CCACAAGAGGTTGGAATGATTTC-3') for amplification and the Jk oligonucleotide (5'-GTAGTCTTCTCAACTCTGTTCAC-3') as a probe. *IL-2* gene PCR, which was provided as a DNA loading control, was performed as described previously<sup>10</sup>.

**Analysis of pMX-DEL<sup>CJ</sup> coding joints.** pMX-DEL<sup>CJ</sup> coding joints were amplified by using oligonucleotides pC (5'-GCACGAAGTCTTGAGACCT-3') and IRES REV5 (5'-CTCGACTAAACACATGTAAAGC-3'). Genomic DNA (300 ng) was used in the original amplification with serial fivefold dilutions. Oligonucleotide IR4 (5'-CCCTTGTGTGAATACGCTTG-3') was used as a probe. Cloning and sequencing of pMX-DEL<sup>CJ</sup> coding joints was performed as described previously<sup>15</sup>. *P* values for both Fig. 4b and Fig. 4c were calculated by Student's *t*-test with Welch's correction for unequal variances.

**Analysis of *Tcrb* coding joints.** Db1-Jb1.1 and Db1-Jb1.2 CJs were amplified by using oligonucleotides Db1 us (5'-CCTTCCTTATCTTCAACTC-3') and Jb1.2 ds (5'-CCTGACTTCCACCCGAGGTT-3') with the following PCR parameters: 92 °C for 1 min 30 s, 55 °C for 2 min 30 s, 72 °C for 1 min. Thymic DNA (300 ng) was used in the original amplification (13 cycles). PCR products were digested with *Bgl*III for 3 h before a second round of amplification (30 cycles) to permit the isolation of DJb CJs without the competing amplification from the germline *Tcrb* locus. Southern blot analyses of the PCR products were performed with oligonucleotide Jb (5'-GTAATCAGAGGAAGGATG-3') as a probe. Thymic DNA isolated from three individual mice for each genotype (*WT*, *Artemis*<sup>-/-</sup> and *Artemis*<sup>-/-</sup>:*H2AX*<sup>-/-</sup>) was analysed. Db1-Jb1.1 CJs from *WT* and *H2AX*<sup>-/-</sup> thymic DNA were cloned and sequenced.

**Flow cytometric analyses.** Flow cytometric analyses were performed on a FACSCaliber (BD Biosciences) with fluorescein isothiocyanate-conjugated anti-CD45R/B220 and allophycocyanin-conjugated IgM and the appropriate isotype control. All antibodies were from BD Biosciences. Cell cycle analyses were performed by assessing DNA content by incubation with Hoechst 33342 dye (Invitrogen) for 1 h at 37 °C before flow cytometric analysis. The percentage of cells in G1 and S + G2 + M were approximated with the Dean-Jett-Fox method in FLOW-JO 8.8.6 software.

- Shull, E. R. *et al.* Differential DNA damage signaling accounts for distinct neural apoptotic responses in ATLD and NBS. *Genes Dev.* **23**, 171–180 (2009).
- Lu, L. *et al.* Actin stress fiber pre-extension in human aortic endothelial cells. *Cell Motil. Cytoskeleton* **65**, 281–294 (2008).
- Langer, E. M. *et al.* Ajuba LIM proteins are snail/slug corepressors required for neural crest development in *Xenopus*. *Dev. Cell* **14**, 424–436 (2008).
- Hegedus, E., Kokai, E., Kotlyar, A., Dombradi, V. & Szabo, G. Separation of 1–23-kb complementary DNA strands by urea-agarose gel electrophoresis. *Nucleic Acids Res.* **37**, e112 (2009).

# ATM damage response and XLF repair factor are functionally redundant in joining DNA breaks

Shan Zha<sup>1\*†</sup>, Chunguang Guo<sup>1\*</sup>, Cristian Boboila<sup>1</sup>, Valentyn Oksenysh<sup>1</sup>, Hwei-Ling Cheng<sup>1</sup>, Yu Zhang<sup>1</sup>, Duane R. Wesemann<sup>1</sup>, Grace Yuen<sup>1</sup>, Harin Patel<sup>1</sup>, Peter H. Goff<sup>1</sup>, Richard L. Dubois<sup>2</sup> & Frederick W. Alt<sup>1</sup>

Classical non-homologous DNA end-joining (NHEJ) is a major mammalian DNA double-strand-break (DSB) repair pathway. Deficiencies for classical NHEJ factors, such as XRCC4, abrogate lymphocyte development, owing to a strict requirement for classical NHEJ to join V(D)J recombination DSB intermediates<sup>1,2</sup>. The XRCC4-like factor (XLF; also called NHEJ1) is mutated in certain immunodeficient human patients and has been implicated in classical NHEJ<sup>3–6</sup>; however, XLF-deficient mice have relatively normal lymphocyte development and their lymphocytes support normal V(D)J recombination<sup>5</sup>. The ataxia telangiectasia-mutated protein (ATM) detects DSBs and activates DSB responses by phosphorylating substrates including histone H2AX<sup>7</sup>. However, ATM deficiency causes only modest V(D)J recombination and lymphocyte developmental defects, and H2AX deficiency does not have a measurable impact on these processes<sup>7–9</sup>. Here we show that XLF, ATM and H2AX all have fundamental roles in processing and joining DNA ends during V(D)J recombination, but that these roles have been masked by unanticipated functional redundancies. Thus, combined deficiency of ATM and XLF nearly blocks mouse lymphocyte development due to an inability to process and join chromosomal V(D)J recombination DSB intermediates. Combined XLF and ATM deficiency also severely impairs classical NHEJ, but not alternative end-joining, during IgH class switch recombination. Redundant ATM and XLF functions in classical NHEJ are mediated by ATM kinase activity and are not required for extra-chromosomal V(D)J recombination, indicating a role for chromatin-associated ATM substrates. Correspondingly, conditional H2AX inactivation in XLF-deficient pro-B lines leads to V(D)J recombination defects associated with marked degradation of unjoined V(D)J ends, revealing that H2AX has a role in this process.

Assembly of immunoglobulin and T-cell-receptor variable region exons is initiated by the RAG1 and RAG2 endonuclease (hereafter referred to as RAG), which generates DNA DSBs between a pair of participating V, D, or J coding segments and flanking recombination signal sequences<sup>10</sup>. V(D)J recombination is completed via joining, respectively, of the two coding segments and two recombination signal sequences by classical NHEJ<sup>2</sup>. Whereas XLF-deficient ( $XLF^{\Delta/\Delta}$ ) embryonic stem cells and mouse embryonic fibroblasts are impaired for V(D)J recombination of extra-chromosomal substrates<sup>5</sup>,  $XLF^{\Delta/\Delta}$  mice are only modestly impaired for lymphocyte development, and  $XLF^{\Delta/\Delta}$  pro-B lines, although having increased sensitivity to ionizing radiation (IR), perform nearly normal V(D)J recombination<sup>5</sup>. Thus, unknown factors may compensate for XLF V(D)J recombination functions in developing lymphocytes<sup>5</sup>. Among the candidates, we considered ATM, which is activated by RAG-generated DSBs<sup>7,8,11</sup>. To elucidate whether ATM has an overlapping V(D)J joining function with XLF, we bred  $XLF^{\Delta/\Delta}$  mice<sup>5</sup> with ATM-deficient ( $Atm^{-/-}$ )<sup>12</sup> mice to generate  $XLF^{\Delta/\Delta}Atm^{-/-}$  mice.  $XLF^{\Delta/\Delta}Atm^{-/-}$  mice were live born but

were significantly smaller than control littermates (Supplementary Fig. 1).

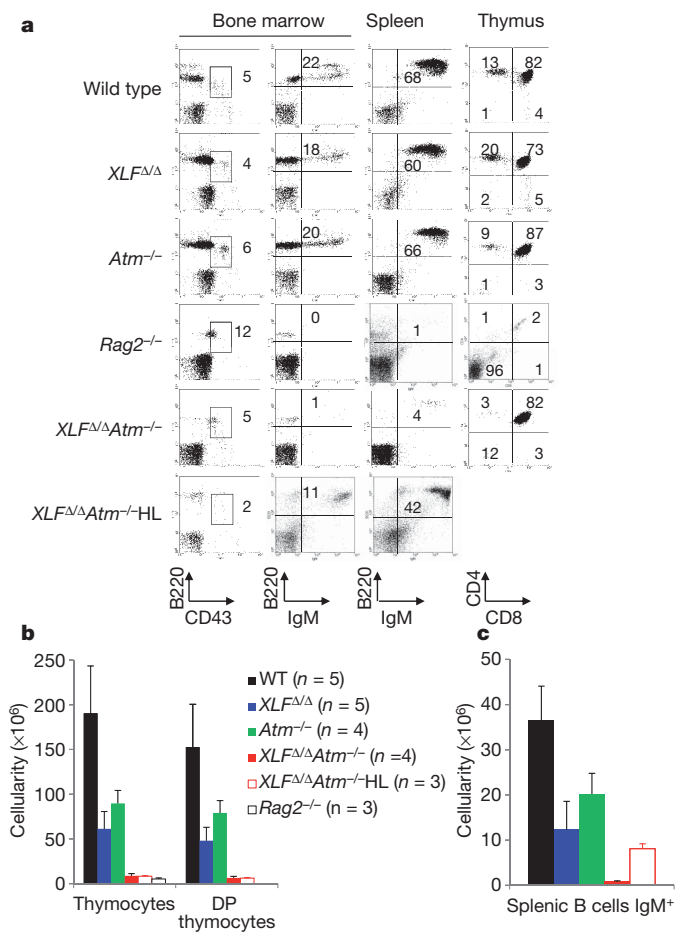
$XLF^{\Delta/\Delta}$  and  $Atm^{-/-}$  mice had only a modest (2–3-fold) reduction in thymocyte numbers and no gross alterations in thymocyte development, as revealed by staining for the CD4 and CD8 differentiation markers (Fig. 1a, b and Supplementary Table 1). In contrast,  $XLF^{\Delta/\Delta}Atm^{-/-}$  mice had a greater than 20-fold decrease in thymocyte numbers, to levels nearly as low as those of  $Rag2^{-/-}$  mice, with an overall developmental pattern reminiscent of that of certain classical NHEJ deficient mice with a 'leaky' V(D)J recombination block<sup>2</sup>. B-cell development was also relatively unimpaired in XLF- and ATM-deficient mice, with both having only modestly reduced (2–3-fold) B220<sup>+</sup>IgM<sup>+</sup> splenic B-cell numbers (Fig. 1a, c and Supplementary Table 1)<sup>5,12</sup>. In contrast,  $XLF^{\Delta/\Delta}Atm^{-/-}$  mice had extremely low splenic B-cell numbers (Fig. 1a, c and Supplementary Table 1). Analyses of bone marrow B-cell development in  $XLF^{\Delta/\Delta}Atm^{-/-}$  mice suggested an impairment at the CD43<sup>+</sup>B220<sup>+</sup> progenitor (pro-) B-cell stage in which V(D)J recombination is initiated, as shown by the near absence of B220<sup>+</sup>CD43<sup>+</sup> precursor B cells (Fig. 1a). To test further whether impaired B-cell development in  $XLF^{\Delta/\Delta}Atm^{-/-}$  mice involved a V(D)J recombination defect, we bred IgH and IgL loci that contained knock-in mutations of pre-assembled IgH and IgL variable region exons (referred to as HL)<sup>13</sup> into the  $XLF^{\Delta/\Delta}Atm^{-/-}$  background and found a significant rescue of B-cell, but not T-cell, development (Fig. 1a–c and Supplementary Table 1). Together, these findings indicate that XLF/ATM double-deficiency severely impairs T- and B-cell development by impairing V(D)J recombination.

To test unequivocally for V(D)J recombination end-joining defects, we generated  $v-abl$  transformed pro-B-cell lines from wild-type,  $XLF^{\Delta/\Delta}$ ,  $Atm^{-/-}$  and  $XLF^{\Delta/\Delta}Atm^{-/-}$  mice that carried *Bcl-2* transgenes<sup>8</sup>. Treatment of  $v-abl$  transformed pro-B lines with STI571, a *v-abl* kinase inhibitor, arrests cells in G1 and induces RAG, leading to efficient V(D)J recombination of integrated substrates in wild-type cells<sup>8</sup>. The *Bcl-2* transgene obviates apoptotic effects of STI571 (ref. 8). We generated multiple pro-B lines from each genotype that harboured, respectively, either a V(D)J recombination substrate designed to assay coding joins (CJs) and unjoined coding ends (CEs) (Fig. 2a) or a substrate designed to assay recombination signal sequence joins (SJs) and unjoined recombination signal sequence ends (SEs) (Fig. 2b). For these experiments, DNA from individual lines was prepared at day 0 (before treatment), day 2 and day 4 of STI571 treatment, digested with restriction endonucleases and assayed for hybridization to the indicated probes (Fig. 2c, d). Wild-type and  $XLF^{\Delta/\Delta}$  lines generated substantial CJ and SJ levels at day 2 and 4 with little or no obvious free CEs, indicative of a classical NHEJ defect (Fig. 2c, d).  $Atm^{-/-}$  lines also generated substantial levels of CJs and SJs, but, consistent with previous studies<sup>8</sup>, also generated a modest level of unjoined CEs at day 2 that appeared partially resolved by day 4 (Fig. 2c). However, there was

<sup>1</sup>Howard Hughes Medical Institute, The Children's Hospital, the Immune Disease Institute and the Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>2</sup>Department of Pathology and Cell Biology, Department of Pediatrics, Institute for Cancer Genetics, Columbia University, New York City, New York 10032, USA. <sup>†</sup>Present address: Department of Pathology and Cell Biology, Department of Pediatrics, Institute for Cancer Genetics, Columbia University, New York City, New York 10032, USA.

\*These authors contributed equally to this work.





**Figure 1 | ATM and XLF have redundant functions in lymphocyte development.** **a**, Representative flow cytometric analyses of bone marrow, spleen and thymus from wild-type,  $XLF^{\Delta/\Delta}$ ,  $Atm^{-/-}$ ,  $Rag2^{-/-}$ ,  $XLF^{\Delta/\Delta}Atm^{-/-}$  and  $XLF^{\Delta/\Delta}Atm^{-/-}HL$  mice (see Methods for further description of mouse lines). Numbers on the plot are percentages of total cells represented by the indicated population. **b**, **c**, Total thymocyte and CD4<sup>+</sup>CD8<sup>+</sup> double positive (DP) thymocyte numbers (**b**) and IgM<sup>+</sup> splenic B-cell numbers (**c**). Each value listed represents the average  $\pm$  standard deviation from at least three mice between 4–12 weeks of age. See Supplementary Table 1 for details.

no obvious recombination signal sequence joining defect in the  $Atm^{-/-}$  lines (Fig. 2d). In contrast,  $XLF^{\Delta/\Delta}Atm^{-/-}$  lines had little accumulation of CJs or SJs at either time point and, instead, accumulated unjoined CE and SEs, respectively (Fig. 2c, d).

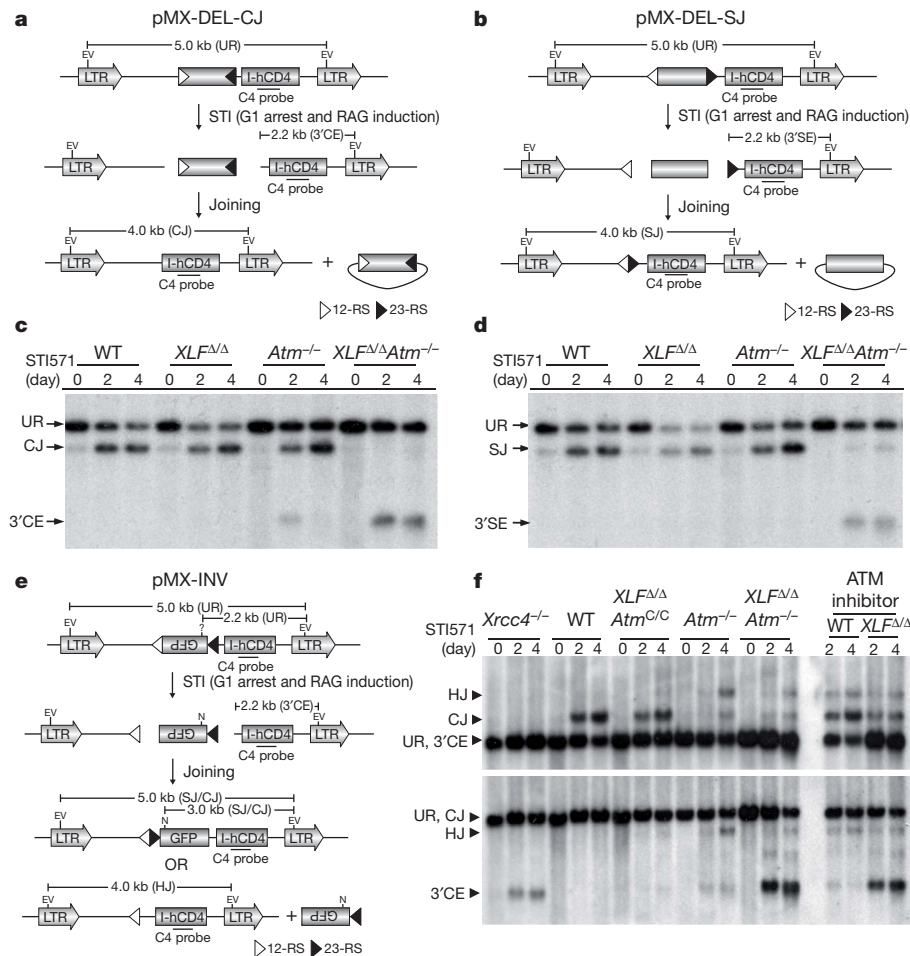
We also tested for V(D)J recombination defects with a substrate that activates a GFP gene upon inversional V(D)J recombination and which, via Southern blotting, reveals CJs, hybrid joins (aberrant joins in which a recombination signal sequence is fused to a coding end) and free CEs (Fig. 2e, f). We clonally integrated a single-copy inversional V(D)J substrate into  $XLF^{\Delta/\Delta}$  pro-B lines that were also homozygous for a conditional knockout ATM allele ( $Atm^{C/C}$ )<sup>12</sup> and then deleted floxed  $Atm^{C/C}$  sequences via Cre recombinase to generate  $XLF^{\Delta/\Delta}Atm^{-/-}$  lines with the same substrate integration. Thus, these matched sets of lines allow assay of a given integrated substrate in  $XLF^{\Delta/\Delta}$  pro-B lines before and after elimination of ATM. We treated inversional substrate containing wild-type,  $Atm^{-/-}$ ,  $XLF^{\Delta/\Delta}Atm^{C/C}$ ,  $XLF^{\Delta/\Delta}Atm^{-/-}$  and  $Xrcc4^{-/-}$  lines with STI571 and assayed for V(D)J recombination both by GFP expression (Supplementary Fig. 2) and Southern blotting (Fig. 2f). Both assays confirmed the severe V(D)J recombination defect in  $XLF^{\Delta/\Delta}Atm^{-/-}$  pro-B lines and Southern blotting confirmed severely defective end-joining, as revealed by a marked decrease in CJs and a marked increase in unjoined CEs (Fig. 2f). The severity of the inversional V(D)J joining defect in  $XLF^{\Delta/\Delta}Atm^{-/-}$  pro-B lines was

similar to that of XRCC4-deficient pro-B lines (Fig. 2f and Supplementary Fig. 2).  $XLF^{\Delta/\Delta}$  pro-B lines treated with an ATM kinase inhibitor also showed a severe end-joining defect during V(D)J recombination, indicating that the ATM-mediated V(D)J joining activity revealed in XLF-deficient lines is mediated by ATM kinase activity (Fig. 2f). Finally, STI571-treated  $XLF^{\Delta/\Delta}Atm^{-/-}$  pro-B lines accumulated unrepaired V(D)J recombination-associated breaks within their endogenous Igk locus, similar to those observed in  $Artemis^{-/-}$  (also called  $Dclre1c$ ) pro-B lines, confirming that the V(D)J recombination defect associated with combined XLF and ATM deficiency extends to this endogenous immunoglobulin locus (Supplementary Fig. 3).

To characterize further the V(D)J recombination defect in  $XLF^{\Delta/\Delta}Atm^{-/-}$  versus wild-type,  $XLF^{\Delta/\Delta}$ ,  $Atm^{-/-}$  and  $Xrcc4^{-/-}$  pro-B lines, we assayed for V(D)J recombination on transiently introduced extra-chromosomal substrates<sup>14</sup>. As this assay is semi-quantitative, within perhaps a fivefold range, and is best for revealing profound defects, we performed at least four independent assays for each genotype (Supplementary Table 1). As expected<sup>2,5</sup>, transient coding and recombination signal sequence joining activity for XRCC4-deficient cells was more than 50-fold less than that of wild-type cells, whereas coding and recombination signal sequence joining activity for  $XLF^{\Delta/\Delta}$  and  $Atm^{-/-}$  cells approached the wild-type range (Supplementary Table 2). Surprisingly, the range of coding and recombination signal sequence joining activity of  $XLF^{\Delta/\Delta}Atm^{-/-}$  pro-B lines, although potentially modestly decreased, overlapped that of wild-type and single mutant cells (Supplementary Table 2). Thus, in contrast to severe defects in chromosomal V(D)J recombination, XLF/ATM double mutant pro-B lines lack severe defects in extra-chromosomal V(D)J recombination.

The chromosomal V(D)J joining defect in  $XLF^{\Delta/\Delta}Atm^{-/-}$  pro-B lines can be attributed to impaired classical NHEJ, as V(D)J recombination exclusively uses this pathway<sup>15</sup>. However, the question remains as to whether  $XLF^{\Delta/\Delta}Atm^{-/-}$  cells have more general classical NHEJ defects and whether they are impaired for other forms of end-joining. IgH class switch recombination (CSR) involves the introduction of DSBs into switch (S) region upstream of the C<sub>μ</sub> constant region exons and their joining to DSBs within a downstream S region, resulting in IgH CSR<sup>16</sup>. Although classical NHEJ is a major CSR joining pathway, CSR occurs at reduced levels in classical NHEJ-deficient cells via alternative end-joining (A-EJ)<sup>13</sup>. To assay CSR, we activated wild-type,  $XLF^{\Delta/\Delta}$ ,  $Atm^{-/-}$  and  $XLF^{\Delta/\Delta}Atm^{-/-}$  HL B cells for 4 days with anti-CD40 and interleukin (IL)-4 to stimulate CSR to IgG1. As expected<sup>15,17–20</sup>,  $XLF^{\Delta/\Delta}$  and  $Atm^{-/-}$  B cells switched to IgG1 at about 40% of wild-type levels<sup>5,17–20</sup>. Moreover,  $XLF^{\Delta/\Delta}Atm^{-/-}$  HL B cells also showed substantial residual IgG1 CSR that was on average about 25% of wild-type levels (Fig. 3a, b and Supplementary Fig. 4). To gain further insight into involved joining pathways, we sequenced the S<sub>μ</sub> to S<sub>γ1</sub> junctions. Classical NHEJ generates CSR junctions with no microhomology (for example, direct joins) and junctions with short (1–2 bp) microhomologies<sup>13</sup>, whereas A-EJ generates CSR junctions that predominantly contain microhomologies<sup>13</sup>. As expected, about 40% of wild-type junctions were direct<sup>13</sup>, whereas only about 22% and 13%, respectively, of  $Atm^{-/-}$  and  $XLF^{\Delta/\Delta}$  junctions were direct<sup>5,20</sup>, indicating some classical NHEJ impairment in these mutant B cells (Fig. 3c and Supplementary Fig. 4). However, only about 5% of  $XLF^{\Delta/\Delta}Atm^{-/-}$  CSR joins were direct (Fig. 3c and Supplementary Fig. 4), consistent with most of their residual CSR being carried out by A-EJ. These results indicate that combined XLF and ATM deficiency impairs general classical NHEJ during CSR, but does not substantially impair A-EJ.

Our finding that the overlapping function of ATM with XLF involves ATM kinase activity and is required for chromosomal versus extra-chromosomal V(D)J joining indicates that this function involves ATM substrates. In response to DSBs, ATM phosphorylates H2AX<sup>7</sup>. However, H2AX deficiency is not known to have a detectable impact on V(D)J recombination<sup>9,21</sup>. To test for overlapping H2AX and XLF



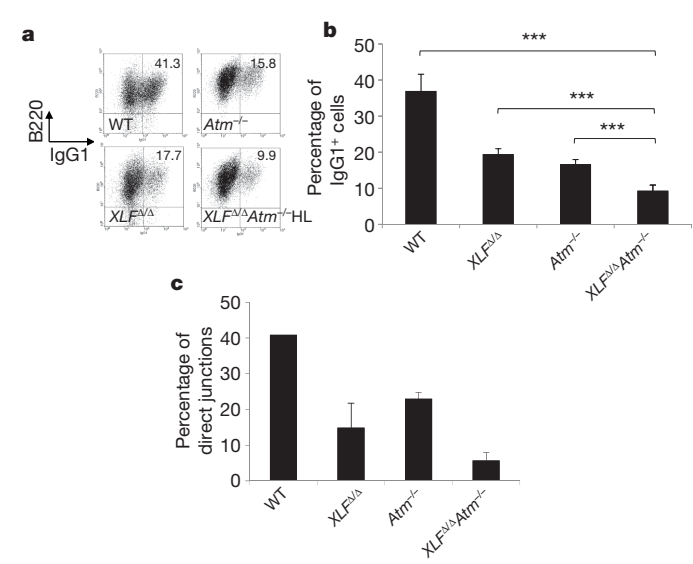
**Figure 2 | ATM and XLF have redundant functions in chromosomal V(D)J recombination.** **a, b, e**, Schematic of pMX-DEL-CJ (**a**), pMX-DEL-SJ (**b**) and pMX-INV (**e**) retroviral recombination substrates designed to assay CJ, SJ and inversionsal V(D)J recombination, respectively<sup>8</sup>. Diagrams indicate un-rearranged substrate (UR), coding/signal end (CE/SE) intermediates and coding joints (CJ/SJ). The 12-recombination signal sequence (12-RS; open triangle), GFP coding sequence, 23-recombination signal sequence (23-RS; filled triangle), IRES-truncated hCD4 cDNA (I-hCD4) and LTRs are indicated. Positions of EcoRV (EV) sites, NcoI (N) sites and C4 probe (black bar) are shown. **c, d**, Southern blotting with C4 probe of EcoRV-digested DNA

functions, we inter-crossed  $XLF^{\Delta/\Delta}$  mice that were heterozygous for an inactivating mutation of H2AX<sup>22</sup> ( $H2AX^{+/F}$ ;  $H2AX$  also called  $H2afx$ ). Notably, these crosses yielded no  $XLF^{\Delta/\Delta}H2AX^{+/F}$  pups, with embryonic death of double homozygous mutants occurring before embryonic day 13.5 (Table 1). The finding that combined XLF and H2AX deficiency, but not combined XLF and ATM deficiency, is embryonically lethal might have several explanations. One is that the lethality results from ATM-independent S-phase functions of H2AX. Another is that impaired checkpoint functions associated with ATM deficiency rescue downstream effects of classical NHEJ deficiency that, otherwise, could result in embryonic death<sup>23</sup>.

To determine whether XLF and H2AX have overlapping V(D)J recombination functions, we generated  $XLF^{\Delta/\Delta}$  mice that were homozygous for a loxP-flanked H2AX allele ( $H2AX^{F/F}$ )<sup>21,22</sup>. From these mice, we generated  $v-abl$  transformed  $XLF^{\Delta/\Delta}H2AX^{F/F}$  pro-B lines containing an integrated single-copy inversional V(D)J recombination substrate (Fig. 4a, b and Supplementary Fig. 5). We then used Cre recombinase to generate matched sets of  $XLF^{\Delta/\Delta}H2AX^{F/F}$  and  $XLF^{\Delta/\Delta}H2AX^{-/-}$  lines, which were treated with STI571 and assayed for V(D)J recombination via GFP expression (Supplementary Fig. 5). In six matched sets, each with a different substrate integration, H2AX

deletion reduced, but did not eliminate, V(D)J recombination (Fig. 4a, b and Supplementary Fig. 5), indicating that XLF and H2AX also have overlapping V(D)J recombination activities, but not to the same extent as XLF and ATM. We also assayed for CJs, hybrid joins and unjoined CEs by Southern blotting in three matched sets of  $XLF^{\Delta/\Delta}H2AX^{F/F}$  and  $XLF^{\Delta/\Delta}H2AX^{-/-}$  pro-B lines (Fig. 4a, b and Supplementary Fig. 4b).  $XLF^{\Delta/\Delta}H2AX^{F/F}$  lines behaved like wild-type or  $XLF^{\Delta/\Delta}$  lines, as they robustly generated CJs but not hybrid joins or unjoined CEs (Fig. 4a, b and Supplementary Fig. 5). However, in accord with GFP assays,  $XLF^{\Delta/\Delta}H2AX^{-/-}$  pro-B lines had substantially reduced CJs compared to  $XLF^{\Delta/\Delta}H2AX^{F/F}$  parents, but did not show readily detectable unjoined CEs (Fig. 4a, b and Supplementary Fig. 5).

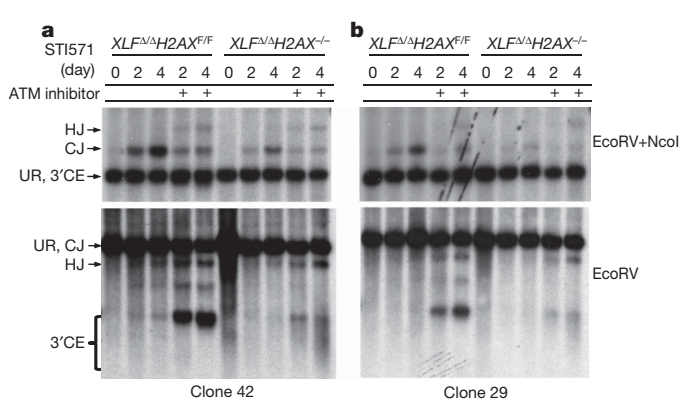
Recent studies found that H2AX protects unjoined coding ends from ATM kinase and CtIP-dependent resection<sup>24</sup>. To test whether reduced V(D)J recombination in  $XLF^{\Delta/\Delta}H2AX^{-/-}$  pro-B lines resulted from reduced RAG cleavage or reduced joining of RAG cleaved ends coupled with end resection, we performed V(D)J joining assays in  $XLF^{\Delta/\Delta}H2AX^{F/F}$  and  $XLF^{\Delta/\Delta}H2AX^{-/-}$  pro-B lines treated with an ATM kinase inhibitor (Fig. 4a, b and Supplementary Fig. 5). ATM inhibitor treatment of the  $XLF^{\Delta/\Delta}H2AX^{F/F}$  lines reproduced the phenotype of  $XLF^{\Delta/\Delta}Atm^{-/-}$  lines, including severely reduced CJs and the



**Figure 3 | ATM and XLF synergize in classical NHEJ during CSR.** **a**, Representative flow cytometric analysis of (at least three independent experiments for each genotype) purified CD43<sup>+</sup> splenocytes from the indicated mice stained for surface B220 and surface IgG1 after a 4-day stimulation with anti-CD40 and IL-4. Additional experiments are in Supplementary Fig. 3. **b**, Summary of IgG1 CSR levels of purified CD43<sup>+</sup> splenocytes after 4 days of anti-CD40 plus IL-4 stimulation. The y-axis shows the average percentage of IgG1<sup>+</sup> cells determined from multiple experiments with cells from wild type (*n* = 5), *XLF*<sup>Δ/Δ</sup> (*n* = 5), *Atm*<sup>-/-</sup> (*n* = 5) and *XLF*<sup>Δ/Δ</sup>*Atm*<sup>-/-</sup> HL (*n* = 8) mice. Error bars show standard deviations. \*\*\**P* < 0.001, based on student's *t*-test between indicated pairs. **c**, Percentage of direct junctions relative to direct plus microhomology (MH)-mediated junctions between Sμ and Sγ1 in anti-CD40-plus IL-4-stimulated B cells. Junctions were obtained from multiple independent experiments with *XLF*<sup>Δ/Δ</sup> (*n* = 4), *Atm*<sup>-/-</sup> (*n* = 3) and *XLF*<sup>Δ/Δ</sup>*Atm*<sup>-/-</sup> HL (*n* = 4) cells. See supplementary Fig. 4 for details.

accumulation of unjoined CE (Figs 2f, 4a, b and Supplementary Fig. 5). However, the ATM-inhibitor-treated *XLF*<sup>Δ/Δ</sup>*H2AX*<sup>-/-</sup> lines now yielded a clear band of unjoined CE associated with a CE smear below the band that is characteristic of aberrant end resection (Fig. 4a, b and Supplementary Fig. 5)<sup>24</sup>. To examine this phenomenon further, we used a sensitive TdT end labelling assay<sup>24</sup>, which revealed unjoined coding ends in STI571-treated *XLF*<sup>Δ/Δ</sup>*H2AX*<sup>-/-</sup> pro-B lines without ATM inhibitor treatment (Supplementary Fig. 6). These results demonstrate that XLF and H2AX have overlapping activities in V(D)J recombinational joining and also indicate that the unjoined V(D)J CE in *XLF*<sup>Δ/Δ</sup>*H2AX*<sup>-/-</sup> pro-B lines are largely resected in the absence of H2AX. Because we did not find complete restoration of the unjoined coding ends in ATM-kinase-inhibitor-treated *XLF*<sup>Δ/Δ</sup>*H2AX*<sup>-/-</sup> pro-B lines, as was observed in ATM-inhibitor-treated classical NHEJ-deficient cells that also are H2AX deficient<sup>24</sup>, XLF might also have an overlapping function with H2AX in end protection.

We consistently observed a lower level of TdT-labelled CE in *XLF*<sup>Δ/Δ</sup> versus *XLF*<sup>Δ/Δ</sup>*H2AX*<sup>-/-</sup> pro-B lines after STI571 induction in the presence of ATM inhibitor, even though the latter have substantially higher levels of unjoined CE (Fig. 4 and Supplementary Figs 5 and 6). This finding indicated that unjoined CE in ATM-inhibitor-treated *XLF*<sup>Δ/Δ</sup> pro-B lines could be blocked from TdT activity. We used urea denaturing gel electrophoresis to test for a defect in opening



**Figure 4 | H2AX and XLF have redundant functions.** **a**, **b**, Southern blot analyses of rearrangement of a clonally integrated single-copy pMX-INV inversional V(D)J recombination substrate in *XLF*<sup>Δ/Δ</sup>*H2AX*<sup>-/-</sup> and derivative *H2AX* deleted *XLF*<sup>Δ/Δ</sup>*H2AX*<sup>-/-</sup> lines. **a**, Clone 42. **b**, Clone 29. Top: DNA was digested with EcoRV and NcoI and probed with C4 probe. Bottom: DNA was digested with EcoRV and probed with C4 probe (see Fig. 2e and legend for additional details). Analysis of an independent line is presented in Supplementary Fig. 5.

coding-end hairpins and found that ATM-inhibitor-treated *XLF*<sup>Δ/Δ</sup> pro-B lines, like *Artemis*<sup>-/-</sup> but not *Xrcc4*<sup>-/-</sup> lines, indeed accumulated unopened hairpin CE (Supplementary Fig. 7). Given the overlapping functions of the ATM kinase and DNA-PK in CSR<sup>25</sup>, plus the role of DNA-PK in activating Artemis to cleave hairpin coding ends<sup>2</sup>, it seemed possible that dual deficiency for ATM and XLF leads to a DNA-PK defect. However, measurements of ionizing-radiation-induced phosphorylation of H2AX and KAP-1—both ATM and DNA-PK substrates<sup>25</sup>—in the presence or absence of DNA-PK inhibitors revealed DNA-PK kinase activity to be as active in *XLF*<sup>Δ/Δ</sup>*Atm*<sup>-/-</sup> cells as in wild-type or *Atm*<sup>-/-</sup> cells (Supplementary Fig. 8). However, we cannot rule out the possibility that *XLF*<sup>Δ/Δ</sup>*Atm*<sup>-/-</sup> cells have a specific defect in DNA-PK activity in the context of V(D)J recombination joining that overlaps with ATM/XLF functions or that a DNA-PK defect in this context could represent a more general defect, for example in the broader recruitment of classical NHEJ factors. In the latter context, we note that the V(D)J recombination defect in *XLF*<sup>Δ/Δ</sup>*Atm*<sup>-/-</sup> pro-B lines is not just in hairpin opening, as we also observe a severe defect in recombination signal sequence joining in these cells (Fig. 2).

ATM, XLF and H2AX previously have been found to have, at most, modest roles in V(D)J recombination and, by extension, classical NHEJ<sup>5,7-9</sup>. Surprisingly, we now find that dual deficiency for XLF and ATM has an impact on V(D)J recombination in progenitor lymphocytes and IgH CSR in mature B cells similarly to deficiency for bona fide classical NHEJ factors. We conclude that XLF-deficient cells require ATM, and that ATM-deficient cells require XLF, to carry out classical NHEJ but not A-EJ. These findings further indicate that XLF-deficient cells provide a novel system for elucidating major, previously unappreciated roles of ATM and ATM substrates in classical NHEJ and vice versa. Our findings already suggest that ATM, via phosphorylation of its substrates, and XLF share overlapping functions primarily in the context of chromosomal classical NHEJ. We have also shown that the V(D)J recombination defects in XLF/ATM- and XLF/H2AX-deficient pro-B lines are not identical either in extent or in outcome. Thus, it remains possible that additional ATM substrates, besides

**Table 1 | H2AX and XLF have redundant functions in murine embryonic development**

Stage	<i>H2AX</i> <sup>+/-</sup> / <i>XLF</i> <sup>Δ/Δ</sup>	<i>H2AX</i> <sup>+/-</sup> / <i>XLF</i> <sup>Δ/Δ</sup>	<i>H2AX</i> <sup>-/-</sup> / <i>XLF</i> <sup>Δ/Δ</sup>	Absorbed	Total
Birth	14	46	0	0	60
Birth (expected)	15	30	15	—	—
E13.5	8	28	0	8	44
E13.5 (expected)	11	22	11	—	—

The indicated genotypes were obtained by inter-crossing *H2AX*<sup>+/-</sup>/*XLF*<sup>Δ/Δ</sup> mice as described<sup>5,12,22</sup>. H2AX deficiency does not cause embryonic lethality<sup>22,27</sup>.



H2AX, may also overlap functionally with XLF. XLF might directly influence the same processes as ATM or H2AX, including end processing and end resection, respectively. Alternatively, overlapping functions might be mediated indirectly through distinct processes. For example, XLF may influence reaction kinetics by classical NHEJ factor recruitment<sup>26</sup>, whereas ATM and ATM substrates seem to tether chromosomal ends for joining<sup>7,8</sup>—two distinct functions that theoretically could be redundant with respect to effects on overall joining activities.

## METHODS SUMMARY

**Mice.** *XLF*<sup>+/Δ</sup>, *Atm*<sup>+/-</sup>, *Atm*<sup>+/-C</sup>, *H2AX*<sup>+/-</sup> and *H2AX*<sup>+/-F</sup> and 'HL' mice have been described previously<sup>5,12,21,22</sup>. All HL mice were heterozygous for both *IgH* and *IgL* knockin alleles.

**Chromosomal V(D)J recombination assays.** V(D)J recombination with an integrated substrate was carried out as described<sup>8</sup>. Briefly, *v-abl* transformed pro-B cell lines were isolated from various mouse lines that harboured an *Eμ-Bcl-2* transgene. For XRCC4-deficient *v-abl* transformants, the *Eμ-Bcl-2* transgene was introduced after establishment of the line<sup>5</sup>. The pro-B lines were infected with the pMX-INV, pMX-DelCJ or pMX-DelSJ retroviral vector and assayed for V(D)J recombination as described<sup>5,8</sup>. ATM inhibitor Ku55933 (Cat.No.118500 from EMD Biosciences) was used as a final concentration of 15 μM as described<sup>8</sup>.

**Lymphocyte development and class switch recombination.** Lymphocyte populations were analysed by flow cytometry as described<sup>5</sup>. Isolation and activation of splenic B cells and flow cytometric assays were as described<sup>13</sup>. Sμ-Sγ1 junctions were isolated from day 4 anti-CD40 plus IL-4 stimulated B cells, cloned and sequenced as described<sup>13</sup>.

Received 17 June; accepted 21 October 2010.

Published online 15 December 2010.

- Lieber, M. R. The mechanism of human nonhomologous DNA end joining. *J. Biol. Chem.* **283**, 1–5 (2008).
- Rooney, S., Chaudhuri, J. & Alt, F. W. The role of the non-homologous end-joining pathway in lymphocyte development. *Immunol. Rev.* **200**, 115–131 (2004).
- Ahnesorg, P., Smith, P. & Jackson, S. P. XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. *Cell* **124**, 301–313 (2006).
- Buck, D. *et al.* Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly. *Cell* **124**, 287–299 (2006).
- Li, G. *et al.* Lymphocyte-specific compensation for XLF/Cernunnos end-joining functions in V(D)J recombination. *Mol. Cell* **31**, 631–640 (2008).
- Zha, S., Alt, F. W., Cheng, H. L., Brush, J. W. & Li, G. Defective DNA repair and increased genomic instability in Cernunnos-XLF-deficient murine ES cells. *Proc. Natl Acad. Sci. USA* **104**, 4518–4523 (2007).
- Bassing, C. H. & Alt, F. W. The cellular response to general and programmed DNA double strand breaks. *DNA Repair (Amst.)* **3**, 781–796 (2004).
- Bredemeyer, A. L. *et al.* ATM stabilizes DNA double-strand-break complexes during V(D)J recombination. *Nature* **442**, 466–470 (2006).
- Yin, B. *et al.* Histone H2AX stabilizes broken DNA strands to suppress chromosome breaks and translocations during V(D)J recombination. *J. Exp. Med.* **206**, 2625–2639 (2009).
- Bassing, C. H., Swat, W. & Alt, F. W. The mechanism and regulation of chromosomal V(D)J recombination. *Cell* **109** (Suppl.), S45–S55 (2002).

- Bredemeyer, A. L. *et al.* DNA double-strand breaks activate a multi-functional genetic program in developing lymphocytes. *Nature* **456**, 819–823 (2008).
- Zha, S., Sekiguchi, J., Brush, J. W., Bassing, C. H. & Alt, F. W. Complementary functions of ATM and H2AX in development and suppression of genomic instability. *Proc. Natl Acad. Sci. USA* **105**, 9302–9306 (2008).
- Yan, C. T. *et al.* IgH class switching and translocations use a robust non-classical end-joining pathway. *Nature* **449**, 478–482 (2007).
- Gellert, M. Molecular analysis of V(D)J recombination. *Annu. Rev. Genet.* **26**, 425–446 (1992).
- Corneo, B. *et al.* Rag mutations reveal robust alternative end joining. *Nature* **449**, 483–486 (2007).
- Chaudhuri, J. *et al.* Evolution of the immunoglobulin heavy chain class switch recombination mechanism. *Adv. Immunol.* **94**, 157–214 (2007).
- Franco, S. *et al.* H2AX prevents DNA breaks from progressing to chromosome breaks and translocations. *Mol. Cell* **21**, 201–214 (2006).
- Ramiro, A. R. *et al.* Role of genomic instability and p53 in AID-induced c-myc-Igh translocations. *Nature* **440**, 105–109 (2006).
- Lumsden, J. M. *et al.* Immunoglobulin class switch recombination is impaired in *Atm*-deficient mice. *J. Exp. Med.* **200**, 1111–1121 (2004).
- Reina-San-Martin, B., Chen, H. T., Nussenzweig, A. & Nussenzweig, M. C. ATM is required for efficient recombination between immunoglobulin switch regions. *J. Exp. Med.* **200**, 1103–1110 (2004).
- Bassing, C. H. *et al.* Increased ionizing radiation sensitivity and genomic instability in the absence of histone H2AX. *Proc. Natl Acad. Sci. USA* **99**, 8173–8178 (2002).
- Bassing, C. H. *et al.* Histone H2AX: a dosage-dependent suppressor of oncogenic translocations and tumors. *Cell* **114**, 359–370 (2003).
- Sekiguchi, J. *et al.* Genetic interactions between ATM and the nonhomologous end-joining factors in genomic stability and development. *Proc. Natl Acad. Sci. USA* **98**, 3243–3248 (2001).
- Helmink, B. A. *et al.* H2AX prevents CtIP-mediated DNA end resection and aberrant repair in G1-phase lymphocytes. *Nature* doi:10.1038/nature09585 (this issue).
- Callén, E. *et al.* Essential role for DNA-PKcs in DNA double-strand break repair and apoptosis in ATM-deficient lymphocytes. *Mol. Cell* **34**, 285–297 (2009).
- Yano, K. *et al.* Ku recruits XLF to DNA double-strand breaks. *EMBO Rep.* **9**, 91–96 (2008).
- Celeste, A. *et al.* Genomic instability in mice lacking histone H2AX. *Science* **296**, 922–927 (2002).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank Y. Fujiwara and P. Huang for technical support. We thank B. Sleckman for advice, reagents and for critical review of this manuscript. This work is supported by NIH grant AI076210 and AI020047 to F.W.A. F.W.A. is an investigator of the Howard Hughes Medical Institute. S.Z. was a fellow, then senior fellow of Leukemia and Lymphomas Society of America and a St Baldrick Scholar. C.G. and Y.Z. are fellows of Cancer Research Institute. C.B. receives support from the pre-doctoral training program of Cancer Research Institute. D.R.W. is supported by a career development award from AAI/GlaxoSmithKline and by NIH training grant AI007376.

**Author Contributions** S.Z., C.G. and F.W.A. designed experiments and wrote the paper. S.Z., C.G., C.B., V.O., H.-L.C., Y.Z., D.R.W., G.Y., H.P., P.H.G. and R.L.D. performed experiments.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to F.W.A. ([alt@enders.tch.harvard.edu](mailto:alt@enders.tch.harvard.edu)).

# CENP-B preserves genome integrity at replication forks paused by retrotransposon LTR

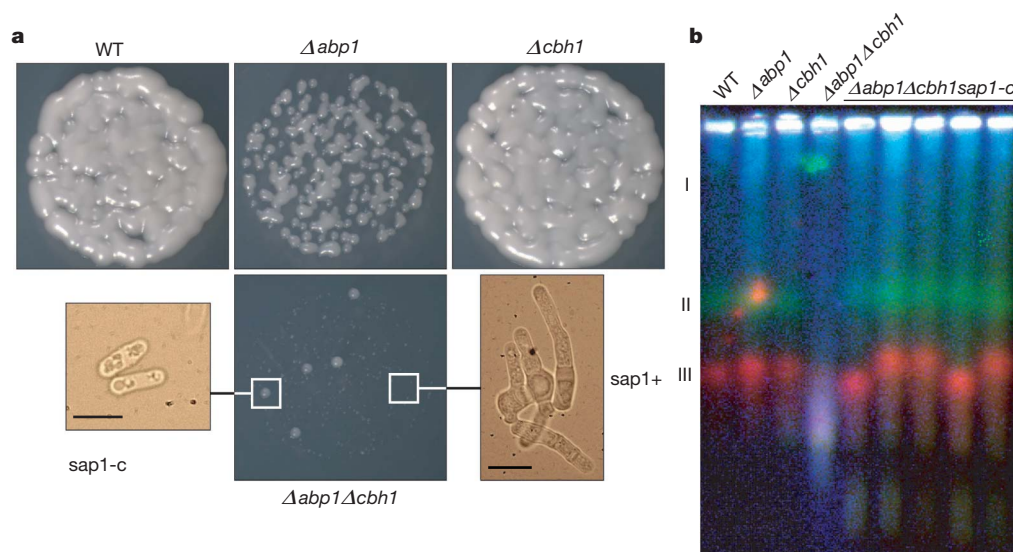
Mikel Zaratiegui<sup>1</sup>, Matthew W. Vaughn<sup>1†</sup>, Danielle V. Irvine<sup>1†</sup>, Derek Goto<sup>1†</sup>, Stephen Watt<sup>2†</sup>, Jürg Bähler<sup>2</sup>, Benoit Arcangioli<sup>3</sup> & Robert A. Martienssen<sup>1</sup>

Centromere-binding protein B (CENP-B) is a widely conserved DNA binding factor associated with heterochromatin and centromeric satellite repeats<sup>1</sup>. In fission yeast, CENP-B homologues have been shown to silence long terminal repeat (LTR) retrotransposons by recruiting histone deacetylases<sup>2</sup>. However, CENP-B factors also have unexplained roles in DNA replication<sup>3,4</sup>. Here we show that a molecular function of CENP-B is to promote replication-fork progression through the LTR. Mutants have increased genomic instability caused by replication-fork blockage that depends on the DNA binding factor switch-activating protein 1 (Sap1), which is directly recruited by the LTR. The loss of Sap1-dependent barrier activity allows the unhindered progression of the replication fork, but results in rearrangements deleterious to the retrotransposon. We conclude that retrotransposons influence replication polarity through recruitment of Sap1 and transposition near replication-fork blocks, whereas CENP-B counteracts this activity and promotes fork stability. Our results may account for the role of LTR in fragile sites, and for the association of CENP-B with pericentromeric heterochromatin and tandem satellite repeats.

In fission yeast, CENP-B proteins are encoded by three homologues, autonomously replicating sequence binding protein 1 (*abp1*), cenp-B homologue 1 (*cbh1*) and *cbh2*, and were previously characterized as DNA binding factors at origins of replication and centromeric repeats, respectively<sup>3,5</sup>. Mutants of *abp1* grow slowly, whereas double mutants with *cbh1* or *cbh2* have severely stunted growth, abnormal mitosis and morphological defects, and triple deletion mutants are inviable<sup>6,7</sup>. As a result, double *Δabp1Δcbh1* mutants form microcolonies on solid

media (Fig. 1a and Supplementary Table 1) and exhibit high levels of cell death (Supplementary Fig. 1). We observed the spontaneous appearance of faster growing cells in a culture of *Δabp1Δcbh1* that grew at rates similar to the *Δabp1* single mutant, lacked morphological defects (Fig. 1a and Supplementary Table 1) and showed lower levels of cell death (Supplementary Fig. 1). Genetic analysis revealed the presence of a single essential locus that also suppressed the lethality of the triple mutant *Δabp1Δcbh1Δcbh2* (not shown). We performed whole-genome resequencing in the mutant strain<sup>8</sup> and isolated a missense mutation in the coding sequence of the DNA binding factor Sap1 (*sap1E101D*, henceforth called *sap1-c*; Supplementary Fig. 2) that co-segregated with suppression of slow growth in *Δabp1Δcbh1* and resulted in lethality in a wild-type background. Sap1 is a protein with essential roles in chromosome stability<sup>9</sup>. Sap1 has been implicated in a programmed replication-fork block in the ribosomal DNA (rDNA) monomer that ensures directional replication to prevent mitotic recombination between rDNA repeats<sup>10–12</sup>.

To test the effects of CENP-B and *sap1-c* mutations on genome integrity, we examined chromosomes by pulsed-field gel electrophoresis. Although single *Δabp1* and *Δcbh1* mutants had wild-type chromosome lengths, the double *Δabp1Δcbh1* mutant had a smear of DNA fragments indicating double-strand breaks in all three chromosomes (Fig. 1b). Treatment of the *Δabp1Δcbh1* sample plugs with the restriction enzyme NotI allowed migration of the chromosomes into the gel, and detection of telomeric and centromeric sequences (Supplementary Fig. 3), suggesting the presence of scattered unresolved replication or recombination intermediates that interfere with the migration of full-length chromosomes,



**Figure 1 | DNA damage in CENP-B mutants is suppressed by *sap1* mutation.** **a**, Images of  $10^3$  plated cells of wild type (WT), *Δabp1*, *Δcbh1* and *Δabp1Δcbh1* with *Δabp1Δcbh1sap1-c* colonies. Microscopy image inserts: branched phenotype in *Δabp1Δcbh1* background (right) and *Δabp1Δcbh1sap1-c* mutant (left). Scale bar, 10  $\mu$ m. **b**, Pulsed-field gel blot analysis of WT, CENP-B mutants (*Δabp1*, *Δcbh1*, *Δabp1Δcbh1*) and five CENP-B/*sap1-c* mutant isolates (*Δabp1Δcbh1sap1-c*). The position of the three chromosomes is indicated on the left. The image is a false-coloured composite of hybridizations for all three chromosomes.

<sup>1</sup>Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, New York 11724, USA. <sup>2</sup>University College London, Department of Genetics, Evolution & Environment, and UCL Cancer Institute, Darwin Building, Gower Street, London WC1E 6BT, UK. <sup>3</sup>Institut Pasteur, Dynamics of the Genome Unit, Departments of Genomes and Genetics and Developmental Biology, CNRS URA2171, F-75015 Paris, France. <sup>†</sup>Present addresses: Texas Advanced Computing Center, University of Texas at Austin, 10100 Burnet Road, Austin, Texas 78758, USA (M.W.V.); Chromosome and Chromatin Research, Murdoch Childrens Research Institute, Department of Paediatrics, University of Melbourne, Royal Children's Hospital, Flemington Road, Parkville 3052, Australia (D.V.I.); Creative Research Initiative Sousei, Hokkaido University, North-21, West-10, Kita-ku, Sapporo, 001-0021, JAPAN (D.G.); Cancer Research UK Cambridge Research Institute, Li Ka Shing Centre, Cambridge CB2 0RE, UK (S.W.).

but not with NotI-digested DNA, into the pulsed field gel. This indicates that Abp1 and Cbh1 have roles in the maintenance of genome integrity.

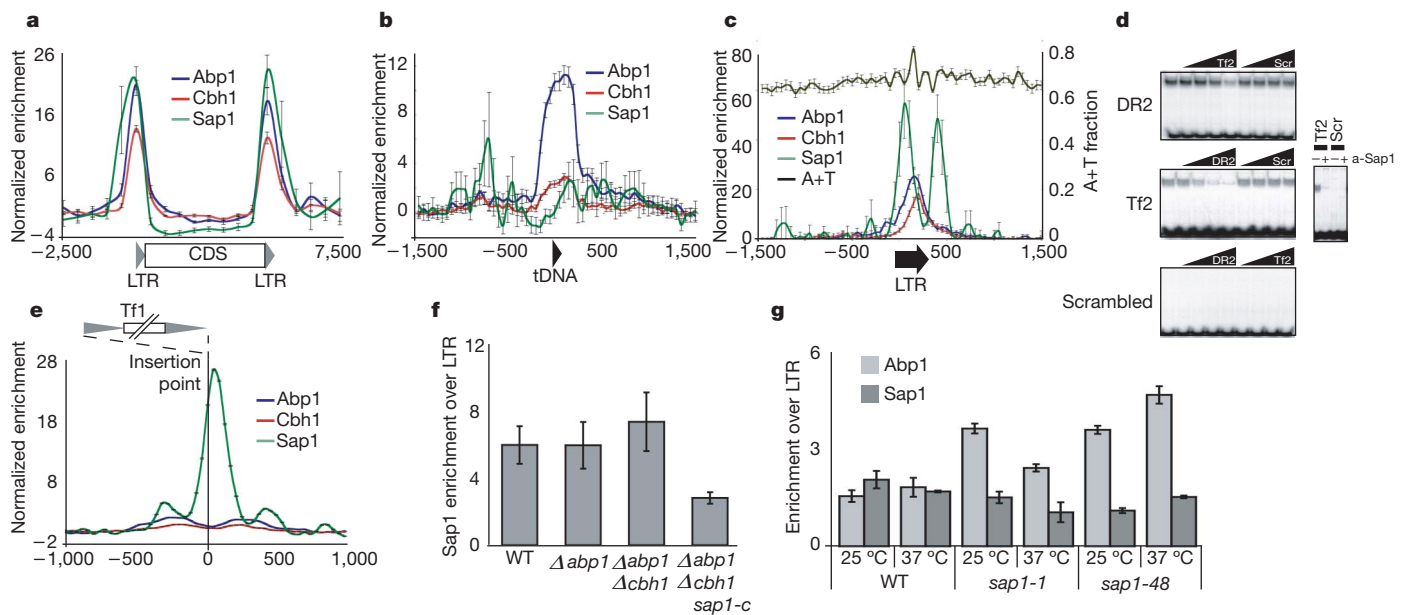
Surprisingly, the *sap1-c* mutation restored genome integrity to all chromosomes, with chromosome 3 exhibiting size variability in several isolates of  $\Delta abp1\Delta cbh1sap1-c$  mutant (Fig. 1b). In fission yeast, chromosome 3 harbours the rDNA repeats. Temperature-sensitive alleles of *sap1* exhibit changes in the size of chromosome 3 attributable to loss of fork barrier activity and an increase in mitotic recombination at rDNA<sup>12</sup>. The changes in the size of chromosome 3 in the *sap1-c* mutants are associated with altered rDNA copy number (Supplementary Fig. 4). The temperature-sensitive alleles *sap1-1* and *sap1-48* (ref. 12) suppressed slow growth in the  $\Delta abp1\Delta cbh1$  double mutant, mimicking *sap1-c* (Supplementary Fig. 5). Consistent with a reduction in fork barrier activity, a probe containing a canonical Sap1-binding sequence had reduced electrophoretic mobility shift in crude extracts from *sap1-c* mutants (Supplementary Fig. 6). We conclude that the suppression of the  $\Delta cbh1\Delta abp1$  phenotype is not specific to the *sap1-c* mutation but a result of defective function of Sap1, and therefore that the loss of genome integrity in  $\Delta abp1\Delta cbh1$  mutants is a consequence of Sap1 activity.

Blocked replication forks are potential sources of genome instability because they can lead to collapse of the replisome and double-strand break formation<sup>13</sup>. The fact that Sap1 activity leads to DNA damage in the absence of Abp1/Cbh1 suggests that the function of CENP-B is to manage Sap1-arrested replication forks. In the absence of Sap1, loss of replication-fork blockage would render Abp1/Cbh1 activity unnecessary and lead to increased genome stability in  $\Delta abp1\Delta cbh1$  mutants. This model predicts that CENP-B and Sap1 would co-localize to the regions where they acted on the replication fork, and that these regions would engage in homologous recombination and degrade to double-strand breaks in the absence of CENP-B. To test this hypothesis, we performed chromatin immunoprecipitation of Sap1, Abp1 and Cbh1 followed by high-throughput sequencing (ChIP-seq). Abp1 has previously been shown to localize and recruit Cbh1 to the LTRs of Tf1 and Tf2 retrotransposons, where Abp1/Cbh1 play a role in their transcriptional silencing<sup>2</sup>. We demonstrated a strong co-localization of Sap1 with Abp1 and Cbh1 at these LTRs as well as at solo LTRs scattered throughout the genome (Fig. 2a, c and Supplementary Fig. 7a, b) and at the mating type locus (Supplementary Fig. 8), where Sap1 and Abp1 have been described to regulate mating-type switching<sup>14,15</sup>. Both Sap1 and Abp1/Cbh1 also

localized to genomic regions independently of each other, suggesting that they do not form a stable complex or mediate their mutual recruitment. In particular, Abp1 exhibited binding to transfer RNA (tRNA) genes (Fig. 2b and Supplementary Fig. 7b), known to be potent replication pause sites<sup>13,16</sup>. Abp1 and Cbh1 co-localize to a highly A/T-rich region located in positions 100–150 of the LTR (Fig. 2c and Supplementary Fig. 7a, b). The localization of Sap1 within the LTR was concentrated in the first 50 base pairs of sequence (Fig. 2c), coinciding with a predicted Sap1-binding site<sup>17</sup> (Supplementary Fig. 7a, c). We tested this sequence by electrophoretic mobility shift assay and detected specific binding in wild-type extracts (Fig. 2d) as well as decreased binding and altered mobility in extracts from  $\Delta abp1\Delta cbh1sap1-c$  mutants (Supplementary Fig. 7d). Interestingly, solo LTR and full-length Tf2 insertions were associated with a prominent peak of Sap1 binding located outside the 3' end of the transposon sequence (Fig. 2c). These observations indicate that Sap1 binding precedes and possibly guides Tf element integration. To test this prediction, we plotted the average enrichment of Sap1, Abp1 and Cbh1 around more than 70,000 *de novo* Tf1 integration sites recently reported<sup>18</sup>. We observed a dramatic association of these integration sites with a peak of Sap1 binding immediately downstream of the insertion site (Fig. 2e and Supplementary Fig. 8) and no appreciable CENP-B enrichment. These results strongly suggest that Sap1-binding sequences determine the targeting and orientation of Tf retroelement transposition.

To evaluate the mutual influence of Sap1 and Abp1/Cbh1 on LTR binding, we performed ChIP analysis of Sap1 in  $\Delta abp1\Delta cbh1$  mutants and of Abp1 in temperature-sensitive *sap1* mutants that affect DNA-binding activity<sup>12</sup>. Sap1 binding to the LTR was unaffected in  $\Delta abp1$  and  $\Delta cbh1$  mutants, and was slightly increased in  $\Delta abp1\Delta cbh1$  double mutants (Fig. 2f), but consistently reduced (twofold) in  $\Delta abp1\Delta cbh1sap1-c$ . Conversely, Abp1 binding to the LTR was increased between two and three times at the permissive temperature in *sap1-1* and *sap1-48* mutants (Fig. 2g). These results indicate that Sap1 and Abp1/Cbh1 bind to the LTR independently of each other and mutually counteract their recruitment, and that the *sap1-c* mutation impairs its binding to the LTR *in vivo* as well as *in vitro* (Supplementary Fig. 4).

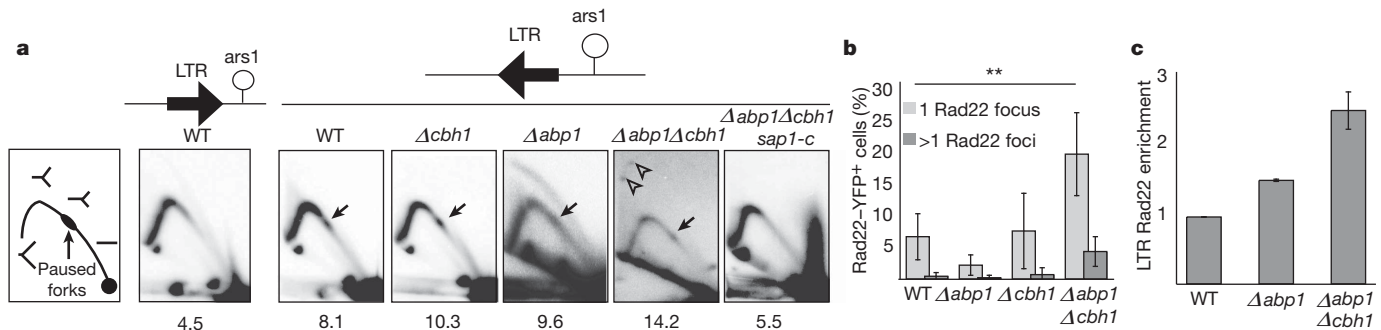
A failure of replication-fork stability at LTRs, which are distributed throughout the genome, would explain the widespread DNA damage in  $\Delta abp1\Delta cbh1$  mutants. We assessed the behaviour of the replication



**Figure 2 | Sap1 and CENP-B co-localize at the LTR of retrotransposons *in vivo*.** Average genome-wide enrichment by ChIP-seq of Sap1, Abp1 and Cbh1 on (a) all Tf2 elements, (b) euchromatic tRNA and (c) solo LTR. Error bars, s.e.m. d, Left panel, competition electrophoretic mobility shift assay; right

panel, inactivation by incubation with anti-Sap1 serum<sup>9</sup>. e, Average Sap1, Abp1 and Cbh1 enrichment around Tf1 *de novo* insertion points<sup>18</sup>. f, ChIP of Sap1 with LTR of Tf2 in CENP-B and *sap1-c* mutants and (g) of Abp1 with LTR of Tf2 in *sap1* temperature-sensitive mutants. Error bars, s.d. for triplicates.



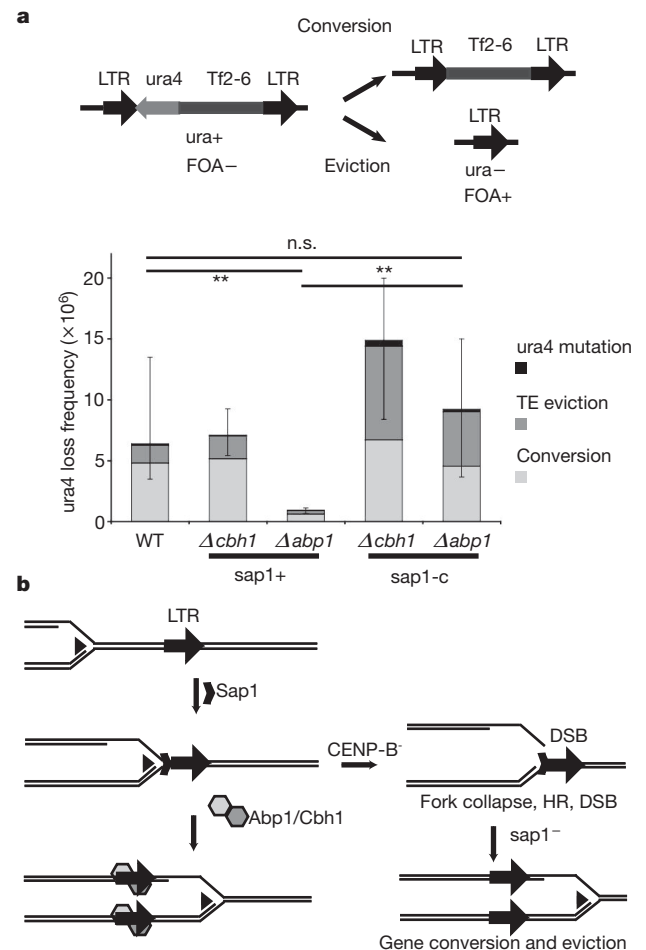


**Figure 3 | CENP-B promotes replication-fork progression through the Sap1-dependent barrier present at the LTR and prevents homologous recombination.** **a**, Two-dimensional gel electrophoresis of a plasmid fragment containing the *Tf2* LTR oriented towards (left) and away (right) from the replication origin *ars1*. Arrows, paused replication intermediates; open arrows, recombination

fork as it traversed the LTR using two-dimensional agarose gel electrophoresis. Sap1-dependent programmed fork blocks are directional and only hinder fork progression in one orientation<sup>10,11</sup>. We cloned a full-length LTR and its first 50 base pairs (containing the Sap1-binding site) in a plasmid in both orientations with respect to the replication origin *ars1*. Two-dimensional gel electrophoresis in a wild-type strain transformed with this episomal system showed a modest accumulation of fork signal at the location of the cloned LTR (Fig. 3a), but only when the Sap1-binding site was proximal to the origin, and not in the opposite orientation (Supplementary Fig. 9). The Sap1-binding site was sufficient for this blocking activity, with the same orientation requirement (Supplementary Fig. 9). We next assayed the LTR for pausing activity in  $\Delta abp1$ ,  $\Delta cbh1$  and  $sap1-c$  mutants (Fig. 3a). Strikingly, the paused fork signal was consistently enhanced and always at the same location in  $\Delta abp1$  and  $\Delta cbh1$  mutants, whereas the  $\Delta abp1\Delta cbh1$  double mutant exhibited additional signals outside the replication arc, suggestive of recombination intermediates<sup>19</sup>. The fork-blocking activity of the LTR disappeared in  $\Delta abp1\Delta cbh1sap1-c$  mutants. Unresolved fork blocks can collapse and undergo homologous recombination for fork recovery. We confirmed the presence of homologous recombination in the  $\Delta abp1\Delta cbh1$  double mutants by measuring the increase in the formation of Rad22 (homologous to Rad52 in *Saccharomyces cerevisiae*) foci in a Rad22–yellow fluorescent protein (YFP) strain<sup>20</sup> (Fig. 3b). We observed that  $\Delta abp1\Delta cbh1$  double-mutant cells accumulated the homologous recombination protein Rad22 at the LTR (Fig. 3c). Consistently, the recombination factor Rhp51 (Rad51 homologue) was essential for viability of  $\Delta abp1\Delta cbh1$  double mutants (Supplementary Fig. 10), indicating that homologous recombination is necessary for recovery from fork stalling at LTRs. These results indicate that Abp1/Cbh1 counteract Sap1 barrier activity and stabilize the replication fork at LTRs. This results in loss of genome integrity and homologous recombination at the LTR in  $\Delta abp1\Delta cbh1$  mutants.

The Sap1-binding sequence is conserved in Tf1 and Tf2 retrotransposon LTRs (Supplementary Fig. 7c), suggesting that it plays a role in the retrotransposon life cycle. We assayed the effect of *sap1* and *abp1/cbh1* on Tf2 stability by measuring the frequencies of loss of a *ura4* reporter transgene inserted in the Tf2-6 transposon<sup>21</sup>. Mutation of *abp1* resulted in a dramatic decrease of Tf2 ectopic recombination, which returned to normal levels when *sap1* was also mutated (Fig. 4). In the presence of *sap1+* there is a preference for gene conversion, which normally constitutes most ectopic recombination events<sup>22</sup>; however, in  $\Delta abp1sap1-c$  and  $\Delta cbh1sap1-c$  mutants the proportion of eviction and conversion events is similar (Fig. 4). Therefore we propose that the LTR recruits Sap1 to control the direction of transposon replication and increase transposon persistence in the genome, perhaps by coordinating lagging strand synthesis, which prevents single-strand annealing from complementary direct repeats (Supplementary Fig. 11a, b). CENP-B counteracts this activity, possibly by promoting replication-fork progression

through the Sap1-dependent barrier. Thus CENP-B and Sap1 promote genome and transposon integrity, respectively, in a ‘tug-of-war’ between transposon and host. Abp1 stimulates fork progression by recruiting the



**Figure 4 | CENP-B and Sap1 have opposite effects on Tf2 stability.** **a**, Ectopic recombination fluctuation assay. Two potential mechanisms of *ura4* loss from the marked *Tf2-6::ura4* are indicated: gene conversion and eviction by LTR recombination. Columns represent total median *ura4* loss frequency in wild type,  $\Delta abp1$ ,  $\Delta cbh1$ ,  $\Delta abp1sap1-c$  and  $\Delta cbh1sap1-c$  mutants; error bars, 95% confidence intervals. Tints indicate distribution of mode of ectopic recombination events in the *ura4*<sup>−</sup> colonies obtained from wild type ( $n = 93$ ),  $\Delta abp1$  ( $n = 88$ ),  $\Delta cbh1$  ( $n = 94$ ),  $\Delta abp1sap1-c$  ( $n = 91$ ) and  $\Delta cbh1sap1-c$  ( $n = 89$ ) mutants. **b**, Model for the interactions between Abp1, Cbh1, Sap1 and the replication fork at the LTR.

fork-restart protein MCM10 (ref. 4), which has primase activity. Additionally the histone deacetylase Mst1, which has roles in replication-fork stability interacts directly with Cbh1 (ref. 23). In *S. cerevisiae* the histone deacetylase Sir2 silences and inhibits recombination in repetitive DNA<sup>24</sup>. CENP-B factors recruit the histone deacetylases Clr3 and Clr6, which perform LTR silencing<sup>2</sup>. The result of these functions would be to preserve genome integrity at LTRs by preventing DNA damage and recombination. This novel role of CENP-B may not be limited to LTR and tDNA, as mutation of the replication-fork blocking factor *reb1*, which is specific to rDNA repeats, also suppresses the slow growth of the *abp1* mutant<sup>25</sup>. Similarly, our ChIP-seq data indicate that Sap1 may also be implicated in the functionality of the replication terminator RTS1 (Supplementary Fig. 8) in collaboration with Rtf1. In this manner, the function and regulation of the Sap1-bound regions is determined by the binding in their vicinity of different factors affecting replication-fork progression.

Because of their repetitive nature, transposons have a close relationship with replication and recombination. For example, the IS608 transposon of *Escherichia coli* is targeted to the lagging strand and always replicated in the same direction<sup>26</sup>. This might prevent recombination between tandemly arranged copies. We have shown that retrotransposons influence DNA replication through recruitment of directional fork blocking factor Sap1 and that activity of CENP-B is required for replication-fork management. Additionally, retrotransposition is targeted to the genomic localization of Sap1. These mechanisms influence the replicative dynamics of the host genome. The genomes of eukaryotes show widespread colonization by retrotransposons, and pericentromeric satellite repeats are often of transposon origin<sup>27</sup>. When such sequences are arranged as tandem repeats, control of replication direction by CENP-B would prevent chromosome breaks and preserve genome integrity. This mechanism accounts for the role of other regulators of fork progression in inter-LTR recombination<sup>28,29</sup>. In contrast, when flanked by LTR in opposite orientations, fragile sites fail to replicate and result in chromosome breaks<sup>13,30</sup>.

## METHODS SUMMARY

ChIP was performed using tagged TAP-Abp1 and TAP-Cbh1 strains with an Anti-Calmodulin Binding Protein antibody (Millipore) and a polyclonal serum against the native Sap1 protein<sup>9</sup>. High-throughput sequencing was performed on an Illumina G2 genome analyser, and analysed for polymorphism detection or statistical analysis of enrichment. Two-dimensional gel electrophoresis was performed as described<sup>11</sup>; see Supplementary Information for construction of the episomal system. Electrophoretic mobility shift assay was performed as described previously<sup>17</sup>.

Received 15 July; accepted 22 October 2010.

Published online 12 December 2010.

- Okada, T. *et al.* CENP-B controls centromere formation depending on the chromatin context. *Cell* **131**, 1287–1300 (2007).
- Cam, H. P., Noma, K., Ebina, H., Levin, H. L. & Grewal, S. I. Host genome surveillance for retrotransposons by transposon-derived proteins. *Nature* **451**, 431–436 (2008).
- Murakami, Y., Huberman, J. A. & Hurwitz, J. Identification, purification, and molecular cloning of autonomously replicating sequence-binding protein 1 from fission yeast *Schizosaccharomyces pombe*. *Proc. Natl Acad. Sci. USA* **93**, 502–507 (1996).
- Locovei, A. M., Spiga, M. G., Tanaka, K., Murakami, Y. & D'Urso, G. The CENP-B homolog, Abp1, interacts with the initiation protein Cdc23 (MCM10) and is required for efficient DNA replication in fission yeast. *Cell Div.* **1**, 27 (2006).
- Lee, J. K., Huberman, J. A. & Hurwitz, J. Purification and characterization of a CENP-B homologue protein that binds to the centromeric K-type repeat DNA of *Schizosaccharomyces pombe*. *Proc. Natl Acad. Sci. USA* **94**, 8427–8432 (1997).
- Baum, M. & Clarke, L. Fission yeast homologs of human CENP-B have redundant functions affecting cell growth and chromosome segregation. *Mol. Cell. Biol.* **20**, 2852–2864 (2000).
- Irelan, J. T., Gutkin, G. I. & Clarke, L. Functional redundancies, distinct localizations and interactions among three fission yeast homologs of centromere protein-B. *Genetics* **157**, 1191–1203 (2001).
- Irvine, D. V. *et al.* Mapping epigenetic mutations in fission yeast using whole-genome next-generation sequencing. *Genome Res.* **19**, 1077–1083 (2009).

- de Lahondes, R., Ribes, V. & Arcangioli, B. Fission yeast Sap1 protein is essential for chromosome stability. *Eukaryot. Cell* **2**, 910–921 (2003).
- Mejia-Ramirez, E., Sanchez-Gorostiaga, A., Krimer, D. B., Schwartzman, J. B. & Hernandez, P. The mating type switch-activating protein Sap1 is required for replication fork arrest at the rRNA genes of fission yeast. *Mol. Cell. Biol.* **25**, 8755–8761 (2005).
- Krings, G. & Bastia, D. Sap1p binds to Ter1 at the ribosomal DNA of *Schizosaccharomyces pombe* and causes polar replication fork arrest. *J. Biol. Chem.* **280**, 39135–39142 (2005).
- Noguchi, C. & Noguchi, E. Sap1 promotes the association of the replication fork protection complex with chromatin and is involved in the replication checkpoint in *Schizosaccharomyces pombe*. *Genetics* **175**, 553–566 (2007).
- Szilard, R. K. *et al.* Systematic identification of fragile sites via genome-wide location analysis of gamma-H2AX. *Nature Struct. Mol. Biol.* **17**, 299–305 (2010).
- Aguilar-Arnal, L., Marsellach, F. X. & Azorin, F. The fission yeast homologue of CENP-B, Abp1, regulates directionality of mating-type switching. *EMBO J.* **27**, 1029–1038 (2008).
- Arcangioli, B. & Klar, A. J. A novel switch-activating site (SAS1) and its cognate binding factor (SAP1) required for efficient mat1 switching in *Schizosaccharomyces pombe*. *EMBO J.* **10**, 3025–3032 (1991).
- Deshpande, A. M. & Newlon, C. S. DNA replication fork pause sites dependent on transcription. *Science* **272**, 1030–1033 (1996).
- Ghazvini, M., Ribes, V. & Arcangioli, B. The essential DNA-binding protein sap1 of *Schizosaccharomyces pombe* contains two independent oligomerization interfaces that dictate the relative orientation of the DNA-binding domain. *Mol. Cell. Biol.* **15**, 4939–4946 (1995).
- Guo, Y. & Levin, H. L. High-throughput sequencing of retrotransposon integration provides a saturated profile of target activity in *Schizosaccharomyces pombe*. *Genome Res.* **20**, 239–248 (2010).
- Segurado, M., Gomez, M. & Antequera, F. Increased recombination intermediates and homologous integration hot spots at DNA replication origins. *Mol. Cell* **10**, 907–916 (2002).
- Meister, P. *et al.* Nuclear factories for signalling and repairing DNA double strand breaks in living fission yeast. *Nucleic Acids Res.* **31**, 5064–5073 (2003).
- Sehgal, A., Lee, C. Y. & Espenshade, P. J. SREBP controls oxygen-dependent mobilization of retrotransposons in fission yeast. *PLoS Genet.* **3**, e131 (2007).
- Kupiec, M. & Petes, T. D. Allelic and ectopic recombination between Ty elements in yeast. *Genetics* **119**, 549–559 (1988).
- Gomez, E. B., Nugent, R. L., Laria, S. & Forsburg, S. L. *Schizosaccharomyces pombe* histone acetyltransferase Mst1 (KAT5) is an essential protein required for damage response and chromosome segregation. *Genetics* **179**, 757–771 (2008).
- Huang, J. & Moazed, D. Association of the RENT complex with nontranscribed and coding regions of rDNA and a regional requirement for the replication fork block protein Fob1 in rDNA silencing. *Genes Dev.* **17**, 2162–2176 (2003).
- Roguev, A. *et al.* Conservation and rewiring of functional modules revealed by an epistasis map in fission yeast. *Science* **322**, 405–410 (2008).
- Ton-Hoang, B. *et al.* Single-stranded DNA transposition is coupled to host replication. *Cell* **142**, 398–408 (2010).
- Wong, L. H. & Choo, K. H. Evolutionary dynamics of transposable elements at the centromere. *Trends Genet.* **20**, 611–616 (2004).
- Putnam, C. D., Hayes, T. K. & Kolodner, R. D. Specific pathways prevent duplication-mediated genome rearrangements. *Nature* **460**, 984–989 (2009).
- Wallis, J. W., Chrebet, G., Brodsky, G., Rolfe, M. & Rothstein, R. A hyper-recombination mutation in *S. cerevisiae* identifies a novel eukaryotic topoisomerase. *Cell* **58**, 409–419 (1989).
- Lemoine, F. J., Degtyareva, N. P., Lobachev, K. & Petes, T. D. Chromosomal translocations in yeast induced by low levels of DNA polymerase  $\alpha$  model for chromosome fragile sites. *Cell* **120**, 587–598 (2005).

Supplementary Information is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank the Martienssen laboratory, V. Aranda, E. Mejia-Ramirez and F. Antequera for technical advice and discussions, and R. Allshire, E. Noguchi, M. O'Connell, P. Espenshade and the National BioResource Project (T. Nakamura, Japan) for strains. This work was supported by National Institutes of Health grant R01GM076396 to R.A.M., Cancer Research UK grant C9546/A6517 to J.B., l'Agence Nationale de la Recherche grant ANR-06-BLAN-0271 to B.A., a National Health and Medical Research Council C.J. Martin Fellowship to D.V.I. and a Postdoctoral Fellowship from the Spanish Ministry of Education to M.Z.

**Author Contributions** M.Z., B.A. and R.A.M. designed the experiments presented and wrote the paper. M.Z. performed and analysed the experiments. M.W.V. provided bioinformatic analysis. D.G. and D.V.I. provided strains. S.W. and J.B. performed additional experiments.

**Author Information** The sequences from the ChIP-seq experiments are deposited in Sequence Read Archive ([www.ncbi.nlm.nih.gov/sra](http://www.ncbi.nlm.nih.gov/sra)) under accession number SRA024710.2. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to R.A.M. ([martiens@cshl.edu](mailto:martiens@cshl.edu)).

# Taxadiene synthase structure and evolution of modular architecture in terpene biosynthesis

Mustafa Köksal<sup>1</sup>, Yinghua Jin<sup>2,3</sup>, Robert M. Coates<sup>2</sup>, Rodney Croteau<sup>4</sup> & David W. Christianson<sup>1</sup>

With more than 55,000 members identified so far in all forms of life, the family of terpene or terpenoid natural products represents the epitome of molecular biodiversity. A well-known and important member of this family is the polycyclic diterpenoid Taxol (paclitaxel), which promotes tubulin polymerization<sup>1</sup> and shows remarkable efficacy in cancer chemotherapy<sup>2</sup>. The first committed step of Taxol biosynthesis in the Pacific yew (*Taxus brevifolia*)<sup>3</sup> is the cyclization of the linear isoprenoid substrate geranylgeranyl diphosphate (GGPP) to form taxa-4(5),11(12)diene<sup>4</sup>, which is catalysed by taxadiene synthase<sup>5</sup>. The full-length form of this diterpene cyclase contains 862 residues, but a roughly 80-residue amino-terminal transit sequence is cleaved on maturation in plastids<sup>6</sup>. We now report the X-ray crystal structure of a truncation variant lacking the transit sequence and an additional 27 residues at the N terminus, hereafter designated TXS. Specifically, we have determined structures of TXS complexed with 13-aza-13,14-dihydrocopalyl diphosphate (1.82 Å resolution) and 2-fluorogeranylgeranyl diphosphate (2.25 Å resolution). The TXS structure reveals a modular assembly of three  $\alpha$ -helical domains. The carboxy-terminal catalytic domain is a class I terpenoid cyclase, which binds and activates substrate GGPP with a three-metal ion cluster. The N-terminal domain and a third 'insertion' domain together adopt the fold of a vestigial class II terpenoid cyclase. A class II cyclase activates the isoprenoid substrate by protonation instead of ionization, and the TXS structure reveals a definitive connection between the two distinct cyclase classes in the evolution of terpenoid biosynthesis.

Although the first structures of C<sub>10</sub> monoterpene<sup>7</sup>, C<sub>15</sub> sesquiterpene<sup>8,9</sup> and C<sub>30</sub> triterpene<sup>10</sup> cyclases appeared several years ago, the structure of the 'missing link' in this series—a C<sub>20</sub> diterpene cyclase—has been unknown until now. Plant diterpene cyclases such as taxadiene synthase are perhaps the most intriguing because they are the largest terpenoid cyclases (800–900 residues) and they are believed to be the most closely related to the ancestral plant terpenoid synthase<sup>11,12</sup>. Triple-domain plant diterpene synthases are believed to have evolved through the fusion of single-domain and double-domain bacterial diterpene cyclases, which in turn evolved from ancient progenitors<sup>13</sup>.

The two distinct classes of terpenoid cyclase have unrelated protein folds and use different substrate activation mechanisms<sup>14–17</sup>. A class I terpenoid cyclase uses a trinuclear metal cluster liganded by conserved motifs DDXXD and (N,D)DXX(S,T)XXE (bold indicates typical metal ligands) to trigger the ionization of the isoprenoid substrate diphosphate group, which generates a carbocation to initiate catalysis. A class II terpenoid cyclase initiates carbocation formation by general acid catalysis, using the 'middle' aspartic acid in a DXDD motif to protonate an isoprenoid double bond or oxirane moiety. Taxadiene synthase lacks a DXDD motif but contains conserved metal-binding motifs and requires Mg<sup>2+</sup> for optimal catalytic activity<sup>5</sup>, indicating that it functions as a class I terpenoid cyclase.

Expression and analysis of N-terminal truncation variants of taxadiene synthase revealed that deletions of 60 or 79 residues yield

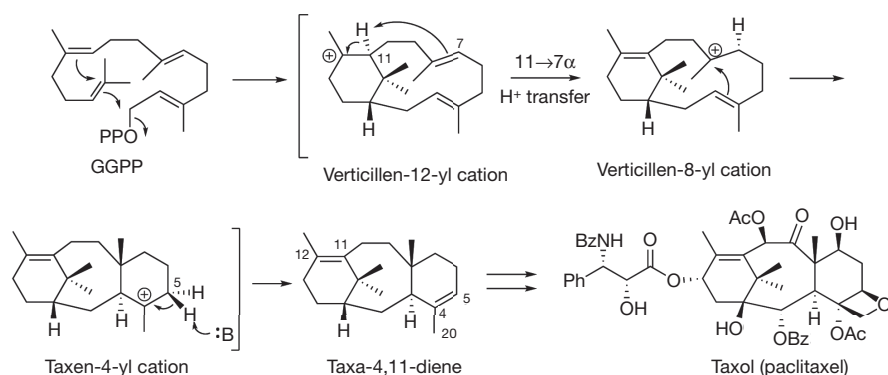
catalytically active proteins, whereas deletions of 93, 113 or 126 residues yield catalytically inactive proteins<sup>18</sup>. These results implicate the N-terminal segment D<sup>80</sup>DIPRLSANYHGDL<sup>93</sup> in catalysis. The N-terminal truncation variant lacking 60 residues has been studied with deuterated<sup>18–21</sup> and fluorinated<sup>22</sup> analogues of GGPP. These studies suggest a cyclization mechanism (Fig. 1) in which the diphosphate leaving group, the 14,15  $\pi$  bond, and the 10,11  $\pi$  bond of GGPP are optimally aligned for leaving-group departure with the formation of a verticillen-12-yl carbocation intermediate in the first step(s) of catalysis. Conformational inversion followed by 11 $\alpha$ ,7 $\alpha$ -proton transfer and transannular B/C ring closure subsequently generates the taxen-4-yl carbocation, whose deprotonation yields taxa-4(5),11(12)-diene. The intramolecular proton transfer required to initiate transannular B/C ring closure occurs without the assistance of an enzyme-bound base<sup>20</sup>. The base mediating the final deprotonation step has not yet been identified.

The successful crystallization of TXS required co-crystallization with Mg<sup>2+</sup> and either 13-aza-13,14-dihydrocopalyl diphosphate (ACP) or 2-fluorogeranylgeranyl diphosphate (FGP) (molecular structures are shown in Supplementary Fig. 1). Although not active, this truncation variant is exceptionally stable and is the only form examined that yielded satisfactory crystals. Surprisingly, TXS contains three  $\alpha$ -helical domains and harbours the folds of both class I and class II terpenoid cyclases (Fig. 2); this structure is representative of nearly all diterpene cyclases. The C-terminal domain (S553–V862) has the class I terpenoid synthase fold that was first observed in farnesyl diphosphate synthase<sup>23</sup> and subsequently observed and designated the class I terpenoid synthase fold<sup>8,14</sup> in monoterpene and sesquiterpene cyclases<sup>7–9</sup>. This fold is also observed in geranylgeranyl diphosphate synthase<sup>24</sup>, which generates the substrate for diterpene cyclases. The N-terminal domain of TXS (M107–I135 and S349–Q552) together with the 'insertion' domain<sup>25</sup> (S136–Y348) comprise the double  $\alpha$ -barrel class II terpenoid synthase fold that was first observed in the triterpene cyclase squalene-hopene cyclase<sup>10</sup> and later observed in oxidosqualene cyclase<sup>26</sup>. TXS shares no significant overall amino-acid sequence identity with these triterpene cyclases.

Comparison of TXS with other terpenoid cyclases reveals that cyclase architecture is modular in nature and can consist of one, two or three domains (Fig. 2). Bacterial and fungal sesquiterpene cyclases are single-domain enzymes that adopt the class I terpenoid synthase fold; the first such enzymes to yield crystal structures were pentalenene synthase<sup>8</sup> and trichodiene synthase<sup>27</sup>, respectively. Plant monoterpene and sesquiterpene cyclases generally contain two domains: the C-terminal domain adopts the class I terpenoid synthase fold, and the N-terminal domain adopts an unrelated  $\alpha$ -helical fold that, as first noted by Wendt and Schulz<sup>14</sup>, is homologous to the N-terminal domain of the class II triterpene cyclase squalene-hopene cyclase<sup>10</sup>. The first plant monoterpene and sesquiterpene synthases to yield crystal structures were bornyl diphosphate synthase<sup>7</sup> and 5-*epi*-aristolochene synthase<sup>9</sup>, respectively. Most plant diterpene synthases contain three

<sup>1</sup>Roy and Diana Vagelos Laboratories, Department of Chemistry, University of Pennsylvania, 231 South 34th Street, Philadelphia, Pennsylvania 19104-6323, USA. <sup>2</sup>Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA. <sup>3</sup>Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309, USA. <sup>4</sup>Institute of Biological Chemistry, Washington State University, Pullman, Washington 99164-6340, USA.





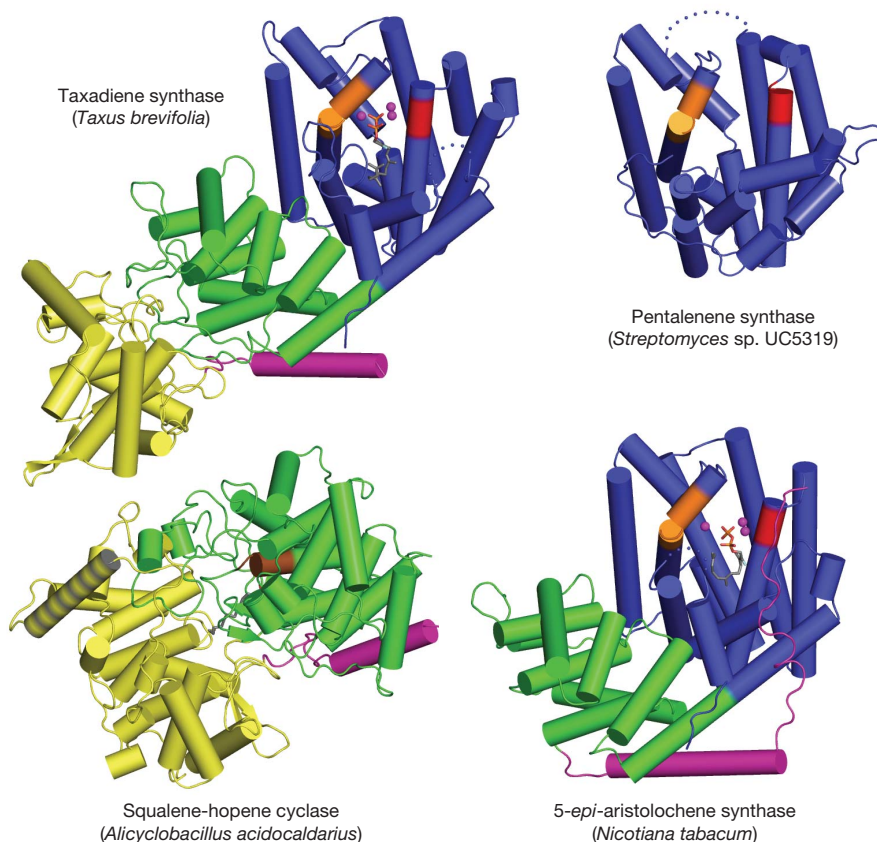
**Figure 1 | Proposed catalytic mechanism of taxadiene synthase.** The cyclization of GGPP to form taxadiene is the first committed step of Taxol (paclitaxel) biosynthesis in yew species. OPP, diphosphate; Ph, phenyl; Ac,

acetyl; Bz, benzyl. Taxadiene is converted to Taxol through a lengthy series of oxidation and acylation steps.

domains, the third being an insertion conserved in sequence and position<sup>25</sup>. It was correctly predicted that this domain is homologous to the insertion domain of a triterpene cyclase on the basis of bioinformatics analysis<sup>13</sup>.

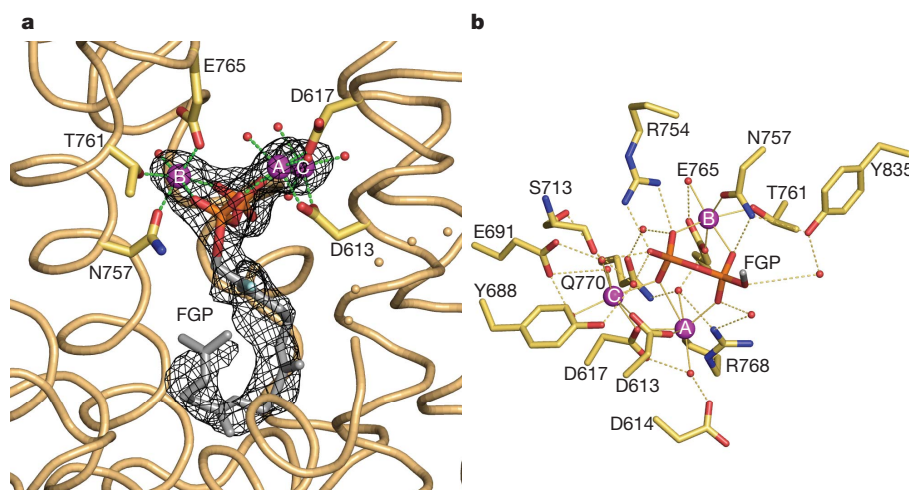
It is interesting to note that the class II triterpene cyclases squalene-hopene cyclase<sup>10</sup> and oxidosqualene cyclase<sup>26</sup> are monotopic membrane proteins: each penetrates, but does not completely pass through, the membranes in which they are localized. Their triterpene substrates

(squalene and squalene oxide, respectively) are solubilized in the membrane and enter the active-site cavity through a hydrophobic channel open to the membrane surface. A nonpolar 'plateau' flanks the entrance to this channel near helix 8 in their respective insertion domains; helix 8 is quite hydrophobic in nature and probably serves as the membrane anchor (Fig. 2). In contrast, TXS functions in the plastid lumen, so its insertion domain does not contain the corresponding hydrophobic components.



**Figure 2 | Structural relationships among terpenoid cyclases.** The class I terpenoid cyclase fold of pentalenene synthase<sup>8</sup> (PDB accession code 1PS1) (blue; 'α domain'<sup>13</sup>) contains metal-binding motifs DDXXD and (N,D)DXX(S,T)XXE (red and orange, respectively); in 5-*epi*-aristolochene synthase<sup>9</sup> (PDB accession code 1LZ9) this domain is linked to a smaller vestigial domain (green; 'β domain'<sup>13</sup>). A related domain is found in the class II terpenoid cyclase fold of squalene-hopene cyclase<sup>10</sup> (PDB accession code

1SQC), where it contains the general acid motif DXDD (brown) and a second domain (yellow; 'γ domain'<sup>13</sup>) inserted between the first (purple) and second helices; a hydrophobic plateau flanking helix 8 (grey stripes) enables membrane insertion. Taxadiene synthase (PDB accession code 3P5R) contains both class I and class II terpenoid cyclase folds, but only the class I domain is catalytically active. The role of N termini (purple) in class I plant cyclases is to 'cap' the active site, as shown for 5-*epi*-aristolochene synthase.



**Figure 3 | Binding of substrate analogue to TXS.** **a**, Simulated annealing  $|F_o| - |F_c|$  omit map in which FGP and three  $Mg^{2+}$  ions are omitted from the structure factor calculation (contoured at  $3.0\sigma$ ); the side chains of metal ligands are indicated. **b**, Molecular recognition of the substrate diphosphate group in the TXS active site. For clarity, the isoprenoid moiety of FGP is truncated to one

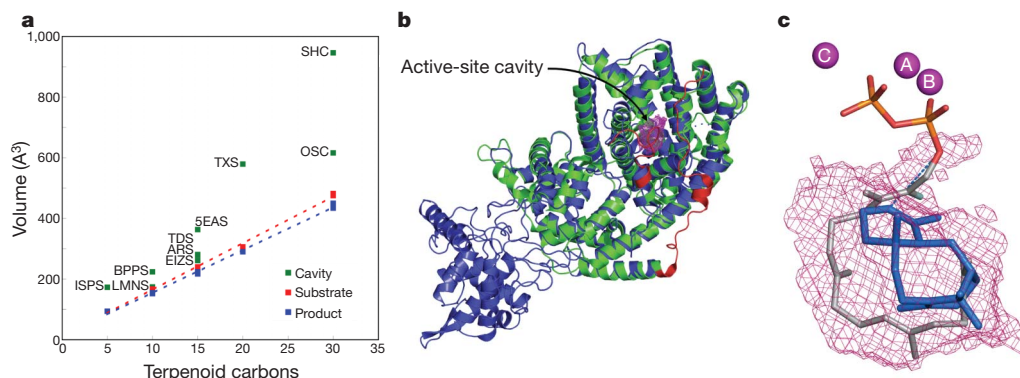
carbon (grey). Metal coordination and hydrogen bond interactions are indicated by thin solid and dashed lines, respectively. Atoms are colour-coded as follows: yellow, carbon; blue, nitrogen; red, oxygen; orange, phosphorus.  $Mg^{2+}$  ions (A, B and C) and water molecules are shown as purple and red spheres, respectively. A corresponding stereo figure is shown in Supplementary Fig. 3.

The active site of TXS is located in the C-terminal domain and is the exclusive binding site of the substrate analogue FGP (Fig. 3a and Supplementary Fig. 3a) and the bicyclic isoprenoid ACP (Supplementary Fig. 2; ACP does not mimic any intermediates in the TXS reaction, although it does mimic a common intermediate of many other diterpene cyclases). Metal-binding motifs that signal class I terpenoid cyclase function<sup>15,27</sup> are conserved in TXS as D<sup>613</sup>DMAD and N<sup>757</sup>DTKTYQAE. The  $Mg^{2+}_A$  and  $Mg^{2+}_C$  ions are coordinated by D613 and D617, and the  $Mg^{2+}_B$  ion is chelated by N757, T761 and E765 (Fig. 3b and Supplementary Fig. 3b). Along with the recent observation of a trinuclear metal cluster in the active site of isoprene synthase<sup>28</sup>, the structure of the TXS- $Mg^{2+}_3$ -FGP complex indicates that three-metal ion catalysis is conserved across the greater family of class I terpenoid synthases: C<sub>5</sub> hemiterpene, C<sub>10</sub> monoterpene, C<sub>15</sub> sesquiterpene and C<sub>20</sub> diterpene synthases.

In addition to metal coordination interactions, the diphosphate group of FGP also accepts hydrogen bonds from R754 and N757 (the latter residue also coordinates to  $Mg^{2+}_B$ ) and makes water-mediated hydrogen bonds with Y688, E691, Y835, S713, R768 and Q770. It is interesting to compare the molecular recognition of the FGP diphosphate group with that of the product diphosphate group in

the plant monoterpene cyclase bornyl diphosphate synthase<sup>7</sup> (Supplementary Fig. 3c). Most residues that assist the trinuclear metal cluster in binding and activating the substrate diphosphate group are conserved between these cyclases.

Class I terpenoid synthases undergo a significant structural transition from an open to a closed active-site conformation after the binding of three  $Mg^{2+}$  ions and the substrate diphosphate group, and this conformational transition helps to protect reactive carbocation intermediates from premature quenching by bulk solvent<sup>15,16</sup>. Although the structure of the fully open conformation of TXS is unavailable, we suggest that the structural changes observed between open and closed active-site conformations in plant monoterpene and sesquiterpene cyclases are representative of those that occur in the plant diterpene cyclase TXS. For example, active-site closure in bornyl diphosphate synthase<sup>7</sup> and 5-*epi*-aristolochene synthase<sup>9</sup> involves conformational changes of loops flanking the mouth of the active site; in addition, the N-terminal polypeptide ‘caps’ each active site. Specifically, the N-terminal polypeptide binds in a groove defined by the A–C and D–D1 loops on one side, and the J–K and H–H- $\alpha$ 1 loops on the other. Tandem arginine residues in the N terminus of bornyl diphosphate synthase make key hydrogen-bond interactions in this groove. The N



**Figure 4 | Active-site cavities of terpenoid synthases.** **a**, Active-site volumes are generally slightly larger than corresponding substrate and product volumes, perhaps to accommodate structural changes better during the cyclization cascade. Abbreviations are defined in Supplementary Table 2. **b**, Superposition of TXS (blue) and bornyl diphosphate synthase (green) guides the modelling of the J–K loop and the N-terminal segment of TXS (red) to define the enclosed active-site cavity (the magenta meshwork indicates the solvent-accessible

surface). **c**, One orientation of taxadiene (blue) fits in the active-site cavity such that the H5 $\beta$  atom of the preceding taxen-4-yl cation would be oriented towards the diphosphate leaving group, suggesting that the PP<sub>i</sub> anion could serve as the stereospecific base that terminates the cyclization cascade. Three  $Mg^{2+}$  ions (A, B and C) and FGP are shown for reference; all protein atoms are omitted for clarity. A corresponding stereo figure is shown in Supplementary Fig. 4.

terminus of 5-*epi*-aristolochene synthase contains only a single corresponding arginine residue, R15, that seems to serve a similar function in the structure of the closed active-site conformation<sup>9</sup>. By analogy with the structures of these plant monoterpene and sesquiterpene cyclases, R84 in the missing N-terminal segment of TXS may help to stabilize the fully closed, catalytically active conformation of mature taxadiene synthase. Accordingly, the closed conformations of the J–K loop and the N-terminal segment of TXS are readily modelled on the basis of the bornyl diphosphate synthase structure to approximate the enclosed active-site contour that serves as the template for GGPP cyclization (Fig. 4 and Supplementary Fig. 4).

The active-site contour of TXS encloses a larger volume than the active sites of monoterpene or sesquiterpene cyclases, which is consistent with the larger isoprenoid substrate of the diterpene cyclase. The active-site cavity volumes of terpenoid synthases correlate with the hydrocarbon volume of their respective isoprenoid substrates (Fig. 4a and Supplementary Table 2). It has been suggested<sup>16</sup> that the shape of the active-site contour is more product-like for high-fidelity cyclases—that is, those that generate a single cyclization product—whereas if the active-site contour is less product-like, a more promiscuous cyclase results that generates multiple cyclization products. For TXS, the active-site volume is significantly larger than the volume of the product taxadiene. This is consistent with the observation that TXS is a somewhat promiscuous cyclase, generating about 20% of the alternative isomer taxa-4(20),11(12)-diene<sup>21</sup>. Indeed, the fact that TXS binds the bicyclic diterpene analogue ACP (Supplementary Fig. 2), which does not correspond to any intermediate in the TXS mechanism, clearly demonstrates promiscuity in ligand binding.

Taxadiene can fit in the enclosed active-site contour of TXS with two alternative orientations (Supplementary Fig. 4). Each orientation leads to possible suggestions for active-site bases that could function in the final deprotonation step of the cyclization cascade. Polar groups in the active site include S587, Q609, Y684, Y688, C719 and C830. Although one of these residues, for example Y688, could conceivably function as a base, taxadiene can also fit within the active-site contour such that H5 $\beta$  of the preceding taxen-4-yl carbocation would be oriented towards the inorganic pyrophosphate (PP<sub>i</sub>) product (Fig. 4c and Supplementary Fig. 4). Thus, the PP<sub>i</sub> anion could serve as a stereospecific base, suggesting the possibility for substrate-assisted or product-assisted catalysis.

Finally, although the N-terminal domain and the insertion domain of TXS form a double  $\alpha$ -barrel class II terpenoid synthase fold such as that characterizing the triterpene cyclases<sup>10,26</sup>, the characteristic general acid DXDD motif and an active-site cavity are absent. Nevertheless, the TXS structure illuminates structure–function relationships in other diterpene cyclases that contain catalytically active class II cyclase domains. For example, consider the bifunctional diterpene cyclase abietadiene synthase from the grand fir tree (*Abies grandis*). Here, the class II terpenoid cyclase domain first catalyses the protonation-dependent cyclization of GGPP to form (+)-copalyl diphosphate, and the class I terpenoid cyclase domain then catalyses the ionization-dependent cyclization of (+)-copalyl diphosphate to form abietadiene<sup>29</sup>. Because the structures of abietadiene synthase and TXS are expected to be homologous, on the basis of 44% amino-acid sequence identity, the protonation-dependent reaction in the class II cyclase domain is presumably catalysed in much the same manner as for a triterpene cyclase reaction. In other diterpene cyclases such as copalyl diphosphate synthase from *Arabidopsis thaliana* (related to TXS by 31% amino-acid sequence identity), only the class II terpenoid cyclase domain is catalytically active; the class I terpenoid cyclase domain is vestigial and the signature metal-binding motifs are absent<sup>30</sup>. Thus, biosynthetic diversity in the family of terpenoid natural products is rooted in a ‘mix and match’ evolutionary strategy with class I and class II terpenoid cyclase folds, which can evolve together or separately as needed to generate the terpenoid product(s) required by the organism.

## METHODS SUMMARY

A variety of different taxadiene synthase constructs were prepared, purified and assessed in crystallization trials, but only one proved satisfactory for crystallization. This construct, designated TXS, was one in which 107 residues were deleted from the N terminus and a hexahistidine tag was added to the C terminus to facilitate purification. TXS was expressed in *Escherichia coli* BL21 (DE3) RIL cells, purified, and co-crystallized with ACP or FGP by the sitting-drop vapour-diffusion method. The initial electron density map of the TXS–ACP complex was phased by using single-wavelength anomalous dispersion. After map fitting, refinement converged smoothly to  $R/R_{\text{free}} = 0.167/0.205$ . The structure of the TXS–FGP complex was solved by molecular replacement and refined to  $R/R_{\text{free}} = 0.187/0.250$ . Data collection and refinement statistics are shown in Supplementary Table 1.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 6 August; accepted 29 October 2010.**

**Published online 15 December 2010.**

- Schiff, P. B., Fant, J. & Horwitz, S. B. Promotion of microtubule assembly *in vitro* by taxol. *Nature* **277**, 665–667 (1979).
- Arbuck, S. G. & Blaylock, B. A. in *Taxol: Science and Applications* (ed. Suffness, M.) 379–415 (CRC Press, 1995).
- Wani, M. C., Taylor, H. L., Wall, M., Coggon, P. & McPhail, A. T. Plant antitumor agents. VI. The isolation and structure of Taxol, a novel antileukemic and antitumor agent from *Taxus brevifolia*. *J. Am. Chem. Soc.* **93**, 2325–2327 (1971).
- Koepp, A. E. *et al.* Cyclization of geranylgeranyl diphosphate to taxa-4(5),11(12)-diene is the committed step of Taxol biosynthesis in Pacific yew. *J. Biol. Chem.* **270**, 8686–8690 (1995).
- Hezari, M., Lewis, N. G. & Croteau, R. Purification and characterization of taxa-4(5), 11(12)-diene synthase from Pacific yew (*Taxus brevifolia*) that catalyzes the first committed step of Taxol biosynthesis. *Arch. Biochem. Biophys.* **322**, 437–444 (1995).
- Wildung, M. R. & Croteau, R. A cDNA clone for taxadiene synthase, the diterpene cyclase that catalyzes the committed step of Taxol biosynthesis. *J. Biol. Chem.* **271**, 9201–9204 (1996).
- Whittington, D. A. *et al.* Bornyl diphosphate synthase: structure and strategy for carbocation manipulation by a terpenoid cyclase. *Proc. Natl Acad. Sci. USA* **99**, 15375–15380 (2002).
- Lesburg, C. A., Zhai, G., Cane, D. E. & Christianson, D. W. Crystal structure of pentalenene synthase: mechanistic insights on terpenoid cyclization reactions in biology. *Science* **277**, 1820–1824 (1997).
- Starks, C. M., Back, K., Chappell, J. & Noel, J. P. Structural basis for cyclic terpene biosynthesis by tobacco 5-*epi*-aristolochene synthase. *Science* **277**, 1815–1820 (1997).
- Wendt, K. U., Poralla, K. & Schulz, G. E. Structure and function of a squalene cyclase. *Science* **277**, 1811–1815 (1997).
- Trapp, S. C. & Croteau, R. Genomic organization of plant terpene synthases and molecular evolutionary implications. *Genetics* **158**, 811–832 (2001).
- Keeling, C. I. *et al.* Identification and functional characterization of monofunctional *ent*-copalyl diphosphate and *ent*-kaurene synthases in white spruce reveal different patterns for diterpene synthase evolution for primary and secondary metabolism in gymnosperms. *Plant Physiol.* **152**, 1197–1208 (2010).
- Cao, R. *et al.* Diterpene cyclases and the nature of the isoprene fold. *Protein Struct. Funct. Bioinf.* **78**, 2417–2432 (2010).
- Wendt, K. U. & Schulz, G. E. Isoprenoid biosynthesis: manifold chemistry catalyzed by similar enzymes. *Structure* **6**, 127–133 (1998).
- Christianson, D. W. Structural biology and chemistry of the terpenoid cyclases. *Chem. Rev.* **106**, 3412–3442 (2006).
- Christianson, D. W. Unearthing the roots of the terpenome. *Curr. Opin. Chem. Biol.* **12**, 141–150 (2008).
- Wendt, K. U., Schulz, G. E., Corey, E. J. & Liu, D. R. Enzyme mechanisms for polycyclic triterpene formation. *Angew. Chem. Int. Ed.* **39**, 2812–2833 (2000).
- Lin, X., Hezari, M., Koepp, A. E., Floss, H. G. & Croteau, R. Mechanism of taxadiene synthase, a diterpene cyclase that catalyzes the first step of Taxol biosynthesis in Pacific yew. *Biochemistry* **35**, 2968–2977 (1996).
- Williams, D. C. *et al.* Heterologous expression and characterization of a ‘pseudomature’ form of taxadiene synthase involved in paclitaxel (Taxol) biosynthesis and evaluation of a potential intermediate and inhibitors of the multistep diterpene cyclization reaction. *Arch. Biochem. Biophys.* **379**, 137–146 (2000).
- Williams, D. C. *et al.* Intramolecular proton transfer in the cyclization of geranylgeranyl diphosphate to the taxadiene precursor of taxol catalyzed by recombinant taxadiene synthase. *Chem. Biol.* **7**, 969–977 (2000).
- Jin, Q., Williams, D. C., Hezari, M., Croteau, R. & Coates, R. M. Stereochemistry of the macrocyclization and elimination steps in taxadiene biosynthesis through deuterium labelling. *J. Org. Chem.* **70**, 4667–4675 (2005).
- Jin, Y., Williams, D. C., Croteau, R. & Coates, R. M. Taxadiene synthase-catalyzed cyclization of 6-fluorogeranylgeranyl diphosphate to 7-fluorovercillenes. *J. Am. Chem. Soc.* **127**, 7834–7842 (2005).
- Tarshis, L. C., Yan, M., Poulter, C. D. & Sacchettini, J. C. Crystal structure of recombinant farnesyl diphosphate synthase at 2.6-Å resolution. *Biochemistry* **33**, 10871–10877 (1994).



24. Chang, T. H., Guo, R. T., Ko, T. P., Wang, A. H. J. & Liang, P. H. Crystal structure of type-III geranylgeranyl pyrophosphate synthase from *Saccharomyces cerevisiae* and the mechanism of product chain length determination. *J. Biol. Chem.* **281**, 14991–15000 (2006).
25. Bohlmann, J., Meyer-Gauen, G. & Croteau, R. Plant terpenoid synthases: molecular biology and phylogenetic analysis. *Proc. Natl Acad. Sci. USA* **95**, 4126–4133 (1998).
26. Thoma, R. *et al.* Insight into steroid scaffold formation from the structure of human oxidosqualene cyclase. *Nature* **432**, 118–122 (2004).
27. Rynkiewicz, M. J., Cane, D. E. & Christianson, D. W. Structure of trichodiene synthase from *Fusarium sporotrichioides* provides mechanistic inferences on the terpene cyclization cascade. *Proc. Natl Acad. Sci. USA* **98**, 13543–13548 (2001).
28. Köksal, M., Zimmer, I., Schnitzler, J.-P. & Christianson, D. W. Structure of isoprene synthase illuminates the chemical mechanism of teragram atmospheric carbon emission. *J. Mol. Biol.* **402**, 363–373 (2010).
29. Peters, R. J., Ravn, M. M., Coates, R. M. & Croteau, R. Bifunctional abietadiene synthase: free diffusive transfer of the (+)-copalyl diphosphate intermediate between two distinct active sites. *J. Am. Chem. Soc.* **123**, 8974–8978 (2001).
30. Prisic, S. & Peters, R. J. Synergistic substrate inhibition of *ent*-copalyl diphosphate synthase: a potential feed-forward inhibition mechanism limiting gibberellin metabolism. *Plant Physiol.* **144**, 445–454 (2007).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank C. MacDermaid and J. Saven for advice and assistance with molecular modelling calculations, and E. Oldfield for helpful comments on the manuscript. We thank the National Synchrotron Light Source at Brookhaven National Laboratory for beamline access. The US National Institutes of Health provided grants GM56838 (D.W.C.), GM13956 (R.M.C.) and CA55254 (R.C.) in support of this research. Y.J. thanks the University of Illinois for support through the John C. Bailar and R. C. Fuson Fellowships.

**Author Contributions** M.K. and D.W.C. performed the X-ray crystallographic studies. R.C. supplied the M79-TXS construct from which M107-TXS-CHT was ultimately prepared. Y.J. and R.M.C. synthesized 2-fluorogeranylgeranyl diphosphate. All authors contributed to the interpretation of the results and preparation of the manuscript.

**Author Information** The atomic coordinates and structure factors of the TXS-ACP and TXS-FGP complexes are deposited in the Protein Data Bank with accession codes 3P5P and 3P5R, respectively. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to D.W.C. ([chris@sas.upenn.edu](mailto:chris@sas.upenn.edu)).

## METHODS

**Cloning, expression and purification of taxadiene synthase.** Heterologous expression of taxadiene synthase from *Taxus brevifolia* lacking the N-terminal segment M1–V79 (M79-TXS) in *Escherichia coli* was achieved at the University of Pennsylvania, using procedures described previously<sup>19</sup>. We found that this protein consistently underwent degradation at 20 °C and 4 °C over a period of a few days to generate a soluble polypeptide stable for at least 4 weeks. Edman sequencing (Wistar Institute Proteomics Facility) showed that this polypeptide lacked the first 29 residues. Given its exceptional stability, this truncated polypeptide was considered a good candidate for crystallization. Accordingly, the M79-TXS gene segment corresponding to an N-terminal truncation at R107 (M107-TXS) was amplified by PCR with the following forward and reverse primers with flanking *NdeI* and *BamHI* sites, respectively: 5'-GCACATATGGAGAGTTCTACTTACCAAGAAC-3' and 5'-GCAGGATCCTACTTGAATTGGATCAATATAAAC-3'. A variant of the pET22b vector (pET22bTV; Novagen) was created by PCR with the following forward and reverse primers with complementary flanking restriction sites: 5'-GCAGGATCCCACCACCACCACCACC-3' and 5'-GCACATATGTATATCTCTCTTAAAGTTAAAC-3'. The gene encoding M107-TXS and the pET22bTV vector were ligated to generate a plasmid encoding the M107-TXS polypeptide with a C-terminal hexahistidine tag (M107-TXS-CHT), which was then used to transform *E. coli* XL1Blue cells for amplification. The resulting clones were confirmed by DNA sequencing (University of Pennsylvania School of Medicine Sequencing Facility) to have only two silent mutations and no amino-acid substitutions.

The M107-TXS-CHT protein (henceforth designated 'TXS') was expressed in *E. coli* BL21 (DE3) RIL cells. Transformed cell cultures were grown in 2-l flasks containing 1 l of Luria–Bertani medium with 100 mg of ampicillin at 37 °C. At an attenuation ( $D_{600}$ ) of 0.6–0.7, cultures were equilibrated at 20 °C and expression was induced by 0.25 mM isopropyl-1-thio- $\beta$ -D-galactopyranoside for 16 h. Cells were harvested by centrifugation at 6,000g for 10 min, producing about 9 g of pellet per litre of culture. The pellet was suspended in 20 ml of buffer E (50 mM  $K_2HPO_4$  pH 7.5, 300 mM NaCl, 10% (v/v) glycerol, 3 mM 2-mercaptoethanol) containing 1 mg ml<sup>-1</sup> lysozyme and 1 mM phenylmethylsulfonyl fluoride, then incubated at 4 °C for 2 h with shaking. Cells were disrupted by sonication on ice six times (30 s on and 90 s off) with a large probe at medium power. Cell debris was cleared by centrifugation twice at 30,000g for 1 h. The clear supernatant was applied to a pre-equilibrated Talon column (Clontech Laboratories) at a flow rate of 1 ml min<sup>-1</sup> with an ÄKTAprime plus fast performance liquid chromatography system (GE Healthcare Bio-Sciences AB). The loaded column was washed three times with 5 column volumes of buffer E, then buffer E plus 5 mM imidazole, then buffer E plus 10 mM imidazole. TXS was eluted with a gradient of 10–200 mM imidazole in buffer E at a flow rate of 2.5 ml min<sup>-1</sup>. Selected fractions were combined, concentrated to a volume of 5 ml, and applied to a Superdex 200 preparatory-grade 26/60 size-exclusion column (GE Healthcare Bio-Sciences AB) with buffer A (25 mM 3-(*N*-morpholino)-2-hydroxypropanesulfonic acid (MOPSO) pH 6.8, 10% (v/v) glycerol, 1 mM dithiothreitol (DTT)) containing 300 mM NaCl. Fractions from this run were combined, concentrated to a volume of 5 ml and applied to the same column a second time with the same buffer. Fractions from the final size-exclusion column were combined and concentrated to 8.6 mg ml<sup>-1</sup>. The purity of the TXS sample was 99% by SDS–PAGE analysis. No hexane-extractable products were identified by gas chromatography–mass spectrometry analysis after incubation with GGPP, indicating that this construct did not generate measurable amounts of taxadiene.

**Crystallization.** TXS could not be crystallized in the absence of isoprenoid diphosphate ligands. However, excellent crystals resulted when the protein was crystallized in the presence of ACP or FGP and  $Mg^{2+}$  ions by the sitting-drop vapour-diffusion method at 4 °C (ligand synthesis is outlined in Supplementary Information). To obtain the crystals of the TXS–ACP complex, a 1- $\mu$ l drop of protein solution (5 mg ml<sup>-1</sup> TXS, 25 mM MOPSO pH 6.8, 10% glycerol, 1 mM DTT, 2.5 mM ACP, 2.5 mM  $MgCl_2$ ) was added to a 1- $\mu$ l drop of precipitant solution (100 mM Bis-Tris pH 6.5, 25% polyethylene glycol 3350, 200 mM NaCl) and equilibrated against a 250- $\mu$ l well reservoir of precipitant solution. Prism-like crystals with rounded edges appeared within 2–3 days and grew to maximal dimensions of 50  $\mu$ m  $\times$  100  $\mu$ m  $\times$  200  $\mu$ m in 2–3 weeks. These crystals were flash-cooled after transfer to a cryoprotectant solution consisting of the mother liquor augmented with 15% ethylene glycol. For the preparation of a heavy-atom derivative for phasing, crystals of the TXS–ACP complex were soaked in a cryoprotectant solution (100 mM HEPES pH 7.5, 25% polyethylene glycol 3350, 100 mM NaCl, 100 mM  $MgCl_2$ , 10% glycerol) containing 2 mM methylmercury chloride for 22 h at 15 °C before flash-cooling.

To obtain crystals of the TXS–FGP complex, a 1- $\mu$ l drop of protein solution (5 mg ml<sup>-1</sup> TXS, 25 mM MOPSO pH 6.8, 10% glycerol, 1 mM DTT, 2.5 mM FGP, 2.5 mM  $MgCl_2$ ) was added to a 1- $\mu$ l drop of precipitant solution (100 mM HEPES pH 7.0, 20% polyethylene glycol 3350, 200 mM  $MgCl_2$ ) and equilibrated against a

250- $\mu$ l well reservoir of precipitant solution. These crystals were flash-cooled after transfer to a cryoprotectant solution consisting of the mother liquor augmented with 10% glycerol.

**Collection and processing of X-ray diffraction data.** Crystals of the TXS–ACP and TXS–FGP complexes diffracted X-rays to 1.82 Å and 2.25 Å resolution, respectively, at the National Synchrotron Light Source (NSLS), Brookhaven National Laboratory, beamline X-29, using incident radiation with  $\lambda = 0.945$  Å and 1.008 Å, respectively. Crystals of the mercury-derivatized TXS–ACP complex diffracted X-rays to 2.6 Å resolution at NSLS beamline X-25 using incident radiation with  $\lambda = 1.000$  Å. All diffraction data were processed with HKL2000 (ref. 31). Crystals of the TXS–ACP complex belonged to space group  $P2_12_12_1$  with unit cell parameters  $a = 55.46$  Å,  $b = 72.41$  Å,  $c = 206.93$  Å, with one molecule in the asymmetric unit; the Matthews coefficient  $V_M$  was  $2.35$  Å<sup>3</sup> Da<sup>-1</sup> (solvent content 48%). Crystals of the TXS–FGP complex belonged to space group  $P2_1$  with unit cell parameters  $a = 54.05$  Å,  $b = 201.98$  Å,  $c = 81.43$  Å,  $\beta = 91.60^\circ$ , with two molecules in the asymmetric unit;  $V_M = 2.61$  Å<sup>3</sup> Da<sup>-1</sup> (solvent content 53%). Data collection and reduction statistics are shown in Supplementary Table 1.

**Phasing and structure refinement.** The initial electron density map of the TXS–ACP complex was phased by single-wavelength anomalous dispersion (SAD) with the 2.6-Å resolution data collected from the methylmercury chloride derivative. Initially, six  $Hg^{2+}$  atoms were located by using the program HKL2MAP<sup>32</sup> and used for SAD phasing, search and refinement of an additional seven  $Hg^{2+}$  sites; density modification, initial electron density map calculation and automatic model building were performed with the AUTOSOL routine implemented in PHENIX<sup>33</sup>. This procedure built more than 50% of the protein residues into the initial electron density map, most of which were  $\alpha$ -helices. Manual model building subsequently generated an initial model with 90% of the residues registered in the sequence. This model was used for molecular replacement calculations with the AUTOMR routine implemented in PHENIX with the 1.82-Å resolution data collected from the TXS–ACP complex. Initial rigid-body refinement, iterative cycles of positional refinement, and grouped and individual atomic  $B$ -factor refinement were performed with PHENIX. Manual model rebuilding was performed with COOT<sup>34</sup>. Water molecules,  $Mg^{2+}$  ions and the ACP molecule were included in later cycles of refinement. A total of 745 out of 764 residues are present in the final model of the TXS–ACP complex; disordered segments excluded from the final model include N-terminal residues M107–S110 (M107 is the N terminus of the construct), the C-terminal hexahistidine tag and its associated linker residues (G863–H870) and surface loop I838–A844. An electron density map of the TXS–ACP complex is shown in Supplementary Fig. 2.

The model of the TXS–ACP complex without its ligand and solvent atoms was used as a search probe for molecular replacement calculations to solve the structure of the TXS–FGP complex at 2.25 Å resolution. Rigid-body refinement, positional refinement and grouped and individual atomic  $B$ -factor refinement were performed with PHENIX. Manual model rebuilding was performed with COOT. In the final model of the TXS–FGP complex, 746 and 736 out of 764 residues were present in monomers A and B, respectively. Disordered segments excluded from the final models of monomers A and B included N-terminal residues M107–S110, the C-terminal hexahistidine tag and its associated linker residues (G863–H870) and loop I574–R578; in addition, surface loop F837–E846 was disordered in monomer B.

For both structures, data reduction and refinement statistics are shown in Supplementary Table 1. Ramachandran plot statistics, calculated with PROCHECK<sup>35</sup>, were as follows. TXS–ACP complex: allowed, 93.7%; additionally allowed, 5.9%; generously allowed, 0.3%; disallowed, 0.1%. TXS–FGP complex: allowed, 91.3%; additionally allowed, 8.1%; generously allowed, 0.4%; disallowed, 0.1%. Simulated-annealing omit maps were calculated with CNS<sup>36</sup>. Protein structure figures were prepared with the graphics program PyMol (<http://www.pymol.org/>).

**Model of TXS in the fully closed conformation.** To model the N terminus and J–K loop segments of TXS in a fully closed conformation and to calculate the active-site cavity volume, the N terminus (residues 54–81) and J–K loop (residues 574–587) segments of bornyl diphosphate synthase in its complex with three  $Mg^{2+}$  ions and 3-azageranyl diphosphate (PDB accession code 1N20) were 'grafted' onto the structure of the TXS- $Mg^{2+}_3$ -FGP complex and mutated to the corresponding residues of TXS; S110 was also introduced to account for an insertion in the sequence alignment between TXS and bornyl diphosphate synthase. The conformations of the grafted segments were then subjected to 10,000 steps of gas-phase conjugate gradient energy minimization, using NAMD<sup>37</sup> and the CHARM22 force field<sup>38</sup>. During energy minimization the grafted segments plus three adjacent residues on the N-terminal and C-terminal ends were unconstrained, while the remaining heavy atoms were fixed. Non-bonded cutoff and switch distances were set to 12 Å and 10 Å, respectively. The final structure resulting from this computation was used as the hypothetical fully closed conformation of TXS. The meshwork representing the active-site cavity of TXS was calculated

with VOIDOO<sup>39</sup>, using a probe with a radius of 1.4 Å to generate a molecular surface based on solvent accessibility. To study product-binding orientations in the enclosed active site, a model of taxadiene was constructed on the basis of the coordinates of the taxane core of Taxol deposited in the Cambridge Crystallographic Data Centre with accession code TEYPAO<sup>40</sup>.

The active-site cavity volume of TXS was compared with the active-site cavity volumes of other terpenoid cyclases, their substrates and their products (Fig. 4a and Supplementary Table 2). All volume calculations were performed with VOIDOO, using a probe with a radius of 0.0 Å to generate a molecular surface based on the atomic van der Waals radii. Because the active site of the hemiterpene synthase isoprene synthase was not fully closed as a result of disorder of the J–K loop, the active-site contour was artificially truncated by the placement of ‘dummy’ atoms to estimate the boundary of the fully enclosed cavity. We used a similar approach to model the cavity of TXS in a chemically sensible manner.

31. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
32. Pape, T. & Schneider, T. R. HKL2MAP: a graphical user interface for phasing with SHELX programs. *J. Appl. Cryst.* **37**, 843–844 (2004).
33. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr.* **D66**, 213–221 (2010).
34. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr.* **D66**, 486–501 (2010).
35. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK—a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291 (1993).
36. Brünger, A. T. *et al.* Crystallography & NMR System: a new software suite for macromolecular structure determination. *Acta Crystallogr.* **D54**, 905–921 (1998).
37. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
38. MacKerell, A. D. Jr *et al.* All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).
39. Kleywegt, G. J. & Jones, T. A. Detection, delineation, measurement and display of cavities in macromolecular structures. *Acta Crystallogr.* **D50**, 178–185 (1994).
40. Gao, Q. & Chen, S. H. An unprecedented side chain conformation of paclitaxel (Taxol®): crystal structure of 7-mesylpaclitaxel. *Tetrahedr. Lett.* **37**, 3425–3428 (1996).



# Genetic variegation of clonal architecture and propagating cells in leukaemia

Kristina Anderson<sup>1</sup>, Christoph Lutz<sup>2</sup>, Frederik W. van Delft<sup>1</sup>, Caroline M. Bateman<sup>1</sup>, Yanping Guo<sup>2</sup>, Susan M. Colman<sup>1</sup>, Helena Kempinski<sup>3</sup>, Anthony V. Moorman<sup>4</sup>, Ian Tittley<sup>1</sup>, John Swansbury<sup>1</sup>, Lyndal Kearney<sup>1</sup>, Tariq Enver<sup>2†</sup> & Mel Greaves<sup>1</sup>

Little is known of the genetic architecture of cancer at the subclonal and single-cell level or in the cells responsible for cancer clone maintenance and propagation. Here we have examined this issue in childhood acute lymphoblastic leukaemia in which the *ETV6-RUNX1* gene fusion is an early or initiating genetic lesion followed by a modest number of recurrent or 'driver' copy number alterations. By multiplexing fluorescence *in situ* hybridization probes for these mutations, up to eight genetic abnormalities can be detected in single cells, a genetic signature of subclones identified and a composite picture of subclonal architecture and putative ancestral trees assembled. Subclones in acute lymphoblastic leukaemia have variegated genetics and complex, nonlinear or branching evolutionary histories. Copy number alterations are independently and reiteratively acquired in subclones of individual patients, and in no preferential order. Clonal architecture is dynamic and is subject to change in the lead-up to a diagnosis and in relapse. Leukaemia propagating cells, assayed by serial transplantation in NOD/SCID IL2R $\gamma^{\text{null}}$  mice, are also genetically variegated, mirroring subclonal patterns, and vary in competitive regenerative capacity *in vivo*. These data have implications for cancer genomics and for the targeted therapy of cancer.

Recent genome-wide scrutiny of cancer cells has revealed extraordinary complexity, with substantial numbers of both potential 'driver' and neutral or 'passenger' mutations per case<sup>1,2</sup>. Informative though these screens are, they probably reflect predominant or composite genetic landscapes that obscure the existence of subclonal heterogeneity of disease<sup>3</sup>. Intracolonial genetic diversity is a common feature of cancer<sup>4</sup> and is probably, from a Darwinian, natural selection perspective, the essential substrate for clonal evolution, disease progression, relapse or metastasis. Subclonal genetic complexity might also be an important consideration for therapeutic targeting. Furthermore, if a subset of 'stem-like' cancer cells, or, as we refer to, propagating cells, are the basis of sustained clonal expansion and disease progression<sup>5</sup> then, in principle, they should be genetically diverse if selection and passage through evolutionary bottlenecks is to occur.

Identifying intracolonial genetic architecture requires genetic scrutiny of single cells or clonal foci, and there are limited examples of this so far<sup>6</sup>; nevertheless, they testify to the existence of significant heterogeneity. The genetic diversity of cancer propagating cells is, as yet, unexplored. We elected to address this issue in lymphoblastic leukaemia. The substantial advantage of this cancer, in addition to its amenability to single-cell analysis, is that it is minimally deranged or unstable, genetically, and the broad, temporal sequence of genetic events is known. For the B-cell precursor subset of childhood acute lymphoblastic leukaemia (ALL) with *ETV6-RUNX1* fusion studied here, the latter genetic lesion is predominantly a prenatal and presumed initiating event<sup>7</sup>. It is coupled with a modest number (3–6) of recurrent, genomic copy number alterations (CNA)<sup>8</sup>. These accrue as secondary and, most likely, postnatal lesions<sup>9</sup> in genes that, predominantly, regulate the cell cycle or B-cell differentiation<sup>8</sup>.

## Subclonal diversity of genotypes in ALL

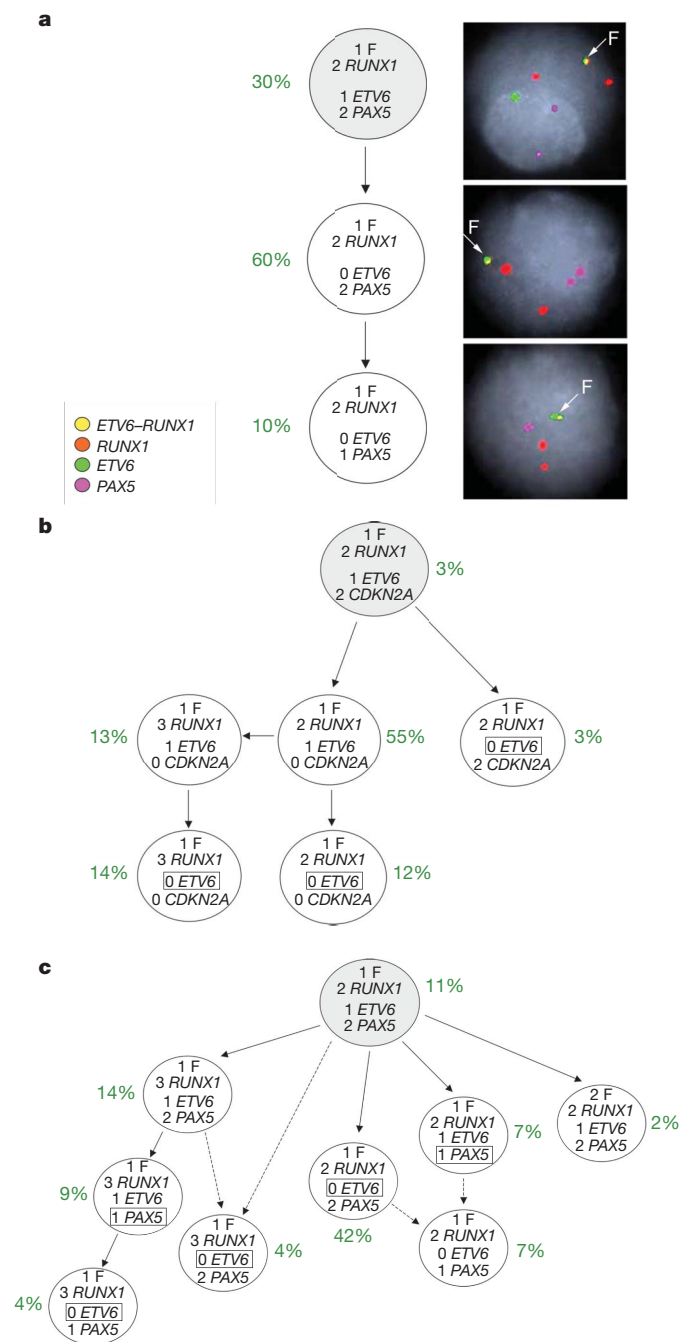
We initially selected 60 cases of *ETV6-RUNX1*-positive ALL and in which *ETV6* was also deleted (15–85% of cells) as detected by fluorescence

*in situ* hybridization (FISH). Of these, 30 were further selected (see Supplementary Table 1) that also had (by FISH) deletion of *PAX5* ( $n = 15$ ) or *CDKN2A* (also called *p16*) ( $n = 12$ ) or deletions of both *PAX5* and *CDKN2A* ( $n = 3$ ) in at least 10% of cells. All 30 cases were then scrutinized using a multiplexed combination of distinctive fluorochrome-labelled bacterial artificial chromosome (BAC) probes. Two-hundred cells with the *ETV6-RUNX1* fusion signal (that is, the reference founder mutation present in all leukaemic cells) were evaluated for each case and each individual cell designed an allele status (that is, mono- or bi-allelic deletion) for *ETV6* and *PAX5* (three colour) or *ETV6* and *CDKN2A* (three colour) or *ETV6*, *PAX5* and *CDKN2A* (four colour). The use of an *ETV6-RUNX1* probe also allowed us to detect duplication of the fusion gene (in 15 out of 30 cases) or an extra copy of chromosome 21q (via *RUNX1* signal copy number; in 21 out of 30 cases). The latter is a common genetic abnormality in ALL and an assumed driver event<sup>10</sup>. Cutoff levels (%) for scoring genetically distinctive subclones were determined using normal blood controls and varied depending upon probe set combination. A threshold was set at 2% for cells with a single CNA (in addition to *ETV6-RUNX1* fusion) and 1% for cells with two or more CNA (see Methods and Supplementary Table 2).

Enumeration of CNA in individual cells in reference to *ETV6-RUNX1* fusion—the universal marker of all the leukaemic cells—allowed us to identify distinctive genetic signatures of subclones and their relative frequencies. From this we could infer the most likely evolutionary or ancestral relationships between the subclones and derive a clonal architecture.

The genetic architectures that were observed were very diverse. The simplest of these genetic architectures that we identified (in 6 cases) involved two or three subclones that could be aligned in a linear sequence (Fig. 1a); however, these were cases with the lowest complement of CNA (only two or three out of the possible total of the seven pre-selected) in addition to the *ETV6-RUNX1* fusion. All

<sup>1</sup>Section of Haemato-Oncology, The Institute of Cancer Research, Sutton SM2 5NG, UK. <sup>2</sup>MRC Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Oxford OX3 9DS, UK. <sup>3</sup>Paediatric Malignancy Unit, Great Ormond Street Hospital & UCL Institute of Child Health, London WC1N 3JH, UK. <sup>4</sup>Leukaemia Research Cytogenetics Group, Northern Institute for Cancer Research, Newcastle University, Newcastle upon Tyne NE1 4LP, UK. <sup>†</sup>Present address: University College London Cancer Institute, London WC1E 6BT, UK.



**Figure 1 | Examples of subclonal architecture in ALL.** **a**, Apparent linear architecture, with three clones (patient no. 13). Representative FISH images on the right show examples of each subclone. **b**, Moderately complex architecture with five subclones (patient no. 8). Loss of the untranslocated *ETV6* allele occurs independently in three separate subclones (boxes). **c**, Complex architecture with eight subclones (patient no. 16). *PAX5* deletions occur independently in two separate subclones (boxes). Arrows indicate probable (or most likely) ancestral derivation of subclones; dashed arrows indicate possible (or alternative) origins of subclones. F, yellow signal, *ETV6*–*RUNX1* fusion gene; 2 *RUNX1*, two red signals (one large, one small) corresponding to one normal *RUNX1* allele, and one small remnant generated from disruption of *RUNX1* allele involved in the gene fusion; *ETV6*, green signal corresponding to the normal (untranslocated) *ETV6* allele; *PAX5* (or *CDKN2A*), pink signal.

other 24 cases had a more marked subclonal heterogeneity with up to ten subclones related via a branching ancestral tree (Fig. 1b, c). Figure 1a–c illustrates examples of the clonal architectures observed (all other cases are depicted in Supplementary Fig. 2).

Inspection of clonal genotypes reveals some previously unrecognized features. It is apparent that the common or highly recurrent CNA are not acquired in any preferential order, indicating that their potency as oncogenic mutations may not be contingent upon (or epistatic to) other CNA. Subclones with the highest number of CNA, positioned ‘terminally’ in the branching architecture, were not necessarily numerically dominant (for example, all three cases illustrated in Fig. 1). Unexpectedly, CNA involving the same gene could be simultaneously present in distinct subclones and must therefore arise more than once, independently. *ETV6* was independently deleted two to three times in 14 of the 30 cases (see Fig. 1b, c), *PAX5* deleted two or three times in 8 out of 18 cases (see Fig. 1c) and *CDKN2A* deleted two times in 4 of 15 cases. This raises interesting mechanistic questions and suggests that these lesions are not only selected on the basis of clonal advantage but may be targeted for DNA-level breakage. One possible mechanism is via off-target effects of RAGS or AID<sup>11–13</sup>.

### Immunophenotypes and genetic diversity

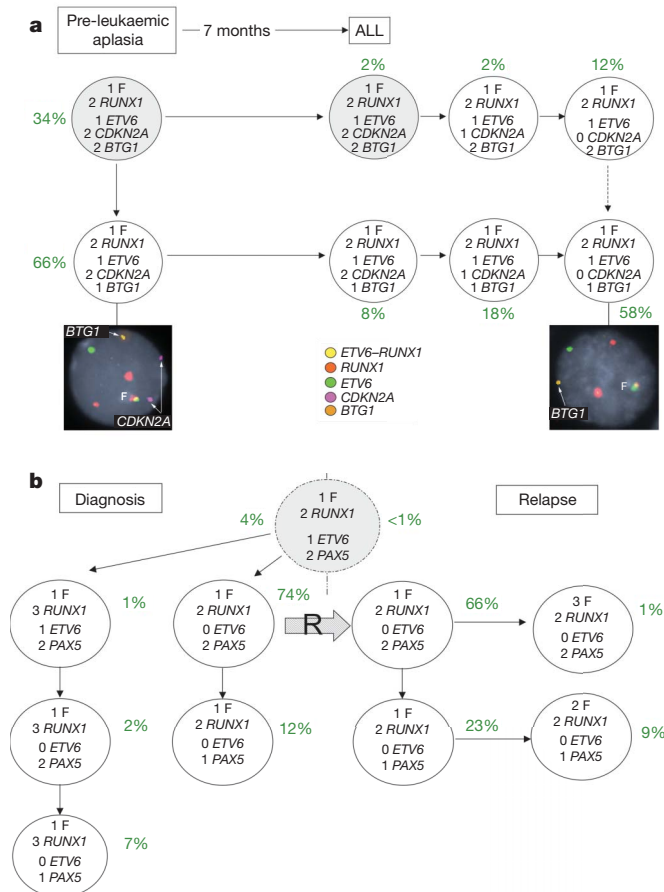
There is a spectrum of early B-lineage differentiation-linked immunophenotypic signatures in ALL<sup>14</sup> and evidence has been presented indicating that cells with several different antibody-defined phenotypes may have leukaemia propagating activity *in vivo*<sup>15</sup>. We analysed the genetic heterogeneity of cells flow sorted on the basis of their expression of CD34 (immature lineage marker) or CD20 (more mature B lineage marker). Sorted populations had similarly complex genetic architectures (Supplementary Fig. 3).

Cells with the immunophenotype CD34<sup>+</sup>CD38<sup>−/low</sup>CD19<sup>+</sup> appear, so far, to be unique to ALL<sup>16</sup>. We previously found this pro-B/stem population, possibly non-activated or quiescent (CD38<sup>−</sup>), to be significantly enriched in ALL propagating cells when assayed in the NOD/SCID strain of mice<sup>17</sup>. When purified by cell sorting, these cells (from patient no. 7) had similar genetic complexity to the bulk leukaemic population (Supplementary Fig. 3).

### Clonal architecture in ALL is dynamic

These descriptions of subclonal, genetic profiles in ALL are snapshots taken at a particular time point, that is, at diagnosis. It is likely that subclonal diversity and the relative dominance of subclones varies continuously with the development and progression of disease. ALL rarely has an identified prodromal phase but occasionally (~2%) patients with ALL have an aplastic, pre-leukaemic phase a few months before a diagnosis of leukaemia<sup>18</sup>. We previously described one such patient with *ETV6*–*RUNX1*<sup>+</sup> ALL<sup>19</sup>. The diagnostic ALL cells had *ETV6*–*RUNX1* fusion but no *ETV6* deletion. Single nucleotide polymorphism (SNP) array screening revealed multiple deletions including *BTG1* and 11q and gain of chromosome X. We compared the clonal genotypes of cells from this patient at these two time points, spread some 7 months apart, and observed a marked shift in clonal architecture (Fig. 2a). The subclones dominating the aplastic, pre-malignant phase were relegated to minor, intermediary subclones in the overt leukaemic phase, with dominance of progeny clones that had homozygously deleted *CDKN2A*. A second patient with a prodromal, aplastic phase some 3 months before a diagnosis of ALL also showed a shift in subclonal dominance (Supplementary Fig. 4).

Treatment and subsequent relapse in ALL reorders the spectrum of genetic abnormalities detected by single gene probing<sup>20</sup> or SNP arrays<sup>21</sup> reflecting the probable selection of distinct subclones as a basis of relapse<sup>21</sup>. For five of the thirty selected patients (numbers 6, 9, 28, 29 and 30), we had matched diagnosis and relapse cells available for multiplexed FISH analysis. Clonal architecture at relapse was different from that at diagnosis. The comparative genetic profiles of subclones allowed us to identify the most likely subclone giving rise to the relapse, although this attribution was not unambiguous (Fig. 2b and Supplementary Fig. 2). Relapse seems to derive from either major or minor clones at diagnosis as previously suggested<sup>21</sup> but with a suggestion that more than one subclone might contribute to relapse



**Figure 2 | Changes in clonal architecture in ALL.** **a**, Scoring for ETV6–RUNX1 fusion and simultaneous deletion of CDKN2A and BTG1 in blast cells at diagnosis of ALL and in bone marrow taken 7 months earlier during a pre-leukaemic aplasia phase. During the pre-leukaemic aplasia phase no deletion of CDKN2A was present, although several other copy number abnormalities were present including 11q, 15q, 5q and BTG1 gene deletions as well as gain of Xq (ref. 19). At diagnosis of ALL some 7 months later, the predominant clone contained homozygous deletion of CDKN2A. Only clones above the cutoffs are shown. Representative, four-colour FISH pictures for the dominant clones during the aplasia and leukaemic phases are shown at the bottom. **b**, An example of relapse originating from a major clone at diagnosis (patient no. 9) (other matched relapse cases for patients 6, 28, 29, 30 are in Supplementary Fig. 2). R, probable clonal origin of relapse.

(for example, patients 6 and 30; Supplementary Fig. 2). The data also indicate that the dominant subclone in relapse itself continues to genetically diversify, in some cases acquiring genetic lesions in the same gene (or chromosome region) as observed in primary, diagnostic subclones. This, along with previous observations on distinctive ETV6 deletions in relapse versus diagnosis<sup>20</sup>, provides further evidence for reiterative CNA. The patterns of genetic diversity observed in relapse indicate that genetically distinct leukaemic propagating cells can survive chemotherapy and provide a reservoir for relapse and further diversification.

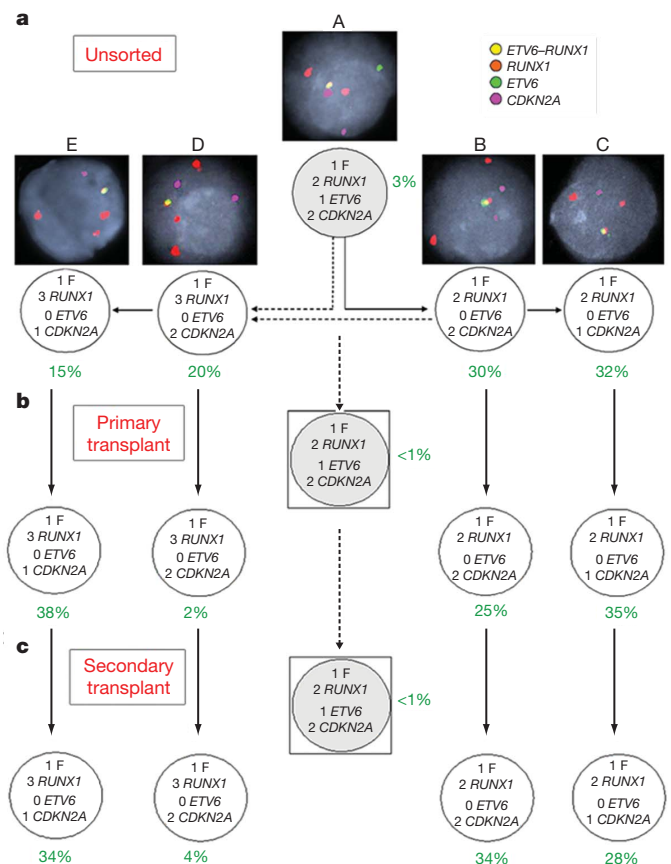
### Genetic diversity of propagating cells in ALL

Within the genetic architecture of ALL, it cannot be assumed that all identified subclones are self-sustaining and propagated by cells with extensive self-renewing capacity<sup>5</sup>. As in evolutionary speciation, it is likely that some branches or subclones are long-lived whereas others are dead ends or out-competed. Nevertheless, the architectural patterns that we observed suggested the possibility that propagating cells for ALL might also have variegated genetics and that this should be demonstrable via serial transplantation in immunodeficient mice. We transplanted, intra-tibially, varying numbers ( $2 \times 10^3$ – $10^6$ ) of unfractionated

or immunophenotypically flow-sorted leukaemic cells into pre-irradiated NOD/SCID IL2R $\gamma^{\text{null}}$  mice. Expanded leukaemic populations were re-transplanted into secondary recipient mice as a validation of self-renewal capacity. We compared the genetic signatures of the cell populations that emerged by successful regeneration *in vivo*, from first and secondary transplants, with those in the original diagnostic sample. Mice with regenerated ALL had significant proportions (3.1 to 93.5 av. 59.4; Supplementary Tables 3 and 4) of human haematopoietic (CD45<sup>+</sup>) cells in the marrow and large, pale spleens (Supplementary Fig. 5b, c). Effectively, all (>99%) human CD45<sup>+</sup> cells were leukaemic with the ETV6–RUNX1 fusion (Supplementary Fig. 5d). Genetic analysis was carried out on cells harvested from bone marrow but when assessed, spleen provided the same result (Supplementary Table 4).

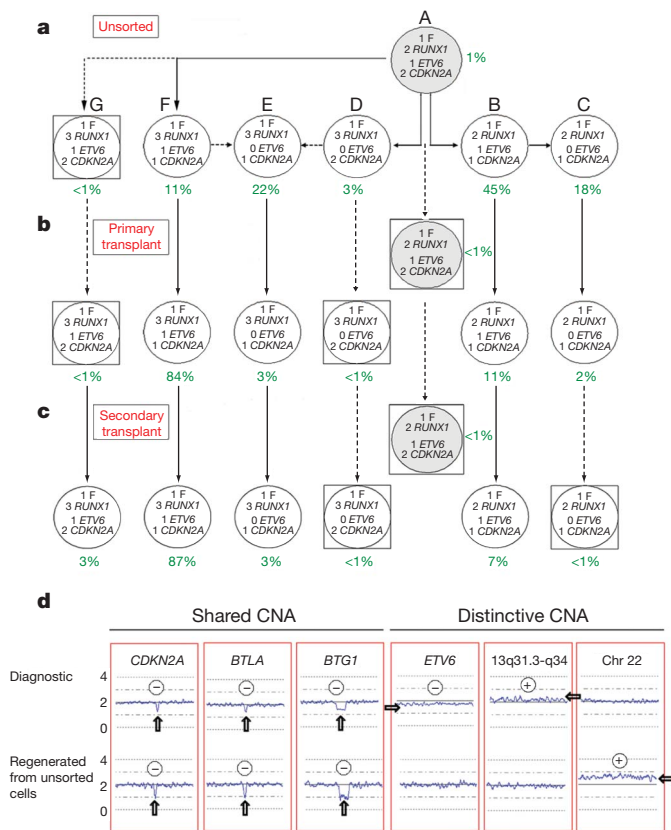
Leukaemic regeneration *in vivo* was observed consistently in both unfractionated populations and in fractions defined immunophenotypically as CD34<sup>+</sup>CD38<sup>−/low</sup>CD19<sup>+</sup> and CD34<sup>+</sup>CD38<sup>+</sup>CD19<sup>+</sup> (Supplementary Tables 3 and 4). This accords with previous evidence<sup>15</sup> that propagating cells in B-cell precursor ALL are not restricted to one immunophenotypic compartment.

In all 24 mice with primary or secondary leukaemic regeneration, several genetically distinct subclones were present, the patterns of which reflected the diversity of subclones identified in the original diagnostic sample (Figs 3 and 4 and Supplementary Tables 3 and 4). The leukaemic cells regenerated in secondary transplants were compared to pre-transplant primary cells by high-resolution SNP arrays. For patient no. 3, these data confirmed subclonal loss of CDKN2A, subclonal gain of chromosome 21 and loss of one copy of ETV6 in



**Figure 3 | Genetics of cells propagating NOD/SCID IL2R $\gamma^{\text{null}}$  mice.** Leukaemic cells from patient no. 3 before injection (**a**), after primary transplantation (**b**; mouse 1, Supplementary Table 3) and after secondary transplantation (**c**; mouse 2, Supplementary Table 3). Representative FISH images are shown of the four subclones (B–E) and the putative pre-leukaemic cell (A) at the top in **a**. Boxed cell, below significance threshold.





**Figure 4 | Shift in clonal architecture of ALL after *in vivo* NOD/SCID IL2R $\gamma$ <sup>null</sup> transplantation.** Clonal architecture of unsorted cells from patient no. 7 before injection (**a**), after primary transplantation (**b**; mouse 1, Supplementary Table 4) and after secondary transplantation (**c**; mouse 2, Supplementary Table 4). **d**, Summary of SNP data on diagnostic versus secondary transplant ALL cells. The blue lines indicate the mean copy number plot of five contiguous SNPs. Middle line, normal diploid copy number. Two fractions of DNA have been examined using 500K SNP arrays: diagnostic DNA (unsorted cells from patient no. 7 before injection) and DNA from leukaemic cells regenerated in mice (mouse 2, Supplementary Table 4). Deletions of *CDKN2A*, *BTLA* and *BTG1* are present in both diagnostic and regenerated samples, whereas deletion of *ETV6*, subclonal gain of the 13q31.3-q34 region and gain of chromosome 22 are distinctive between the two samples. Plus and minus symbols indicate gains or losses of genetic regions, respectively. Arrows highlight CNA.

both diagnostic and regenerated samples (Supplementary Fig. 6 and Supplementary Table 3).

The genetic profiles of regenerated leukaemias revealed variable potency of genetically distinct subclones. For patient no. 3 (Fig. 3 and Supplementary Table 3), four subclones read-out in primary and secondary transplants with three being co-dominant. The putative pre-leukaemic clone (with *ETV6*–*RUNX1* fusion only) did not regenerate. With unfractionated cells from patient no. 7 (Fig. 4 and Supplementary Table 4), four of six subclones regenerated upon secondary transplantation. One of these (subclone F in Fig. 4) was dominant despite being a minor (11%) subclone in the initial diagnostic sample. SNP arrays were used to compare the primary diagnostic sample (patient no. 7) versus the regenerated (secondary) transplant leukaemias. These revealed that both leukaemic populations had *BTLA* and *BTG1* deletions (Fig. 4d) in addition to the CNA in *CDKN2A*, *ETV6* and chromosome 21. The regenerated leukaemia had an additional chromosome 22 that appeared to be absent from the initial, diagnostic cell population (Fig. 4d). This was further investigated by FISH using a chromosome-22-specific BAC probe. This confirmed the SNP array data with the majority of cells (with three *RUNX1* and one *ETV6* signals) in the regenerated leukaemia having an extra chromosome 22 signal (Supplementary Fig. 7). No cells

with an extra chromosome 22 signal were detectable in the initial diagnostic sample in accord with the SNP array data. However, because clone F was dominant in all 9 of 9 mice transplanted with the same cell population (Supplementary Table 4), we assume that a minor subclone of clone F with the extra chromosome 22 was present in the diagnostic sample but at a <1% frequency. Variable competitive potency of subclonal regeneration was also seen in mice injected with immunologically fractionated cells (Supplementary Table 4 and Supplementary Figs 7 and 8).

These data are indicative of additional genetic complexity of subclones and their propagating cells. Moreover, they indicate that distinctive genotypes are associated, functionally, with variable competitive regeneration *in vivo*.

## Discussion

Cancer development at the cellular level is widely regarded as a Darwinian evolutionary process involving ‘natural selection’ of genetically variant cells in the context of a complex micro-environmental ecology<sup>22–24</sup>. Mutational and phenotypic diversity between cells is, in principle, fundamental to this process. Moreover, driver mutations can be expected to have maximal selective currency when present in cells with self-renewing functionality.

Evidence for intracolon genetic diversity in cancer has been provided by chromosome karyotype<sup>25</sup>, by genetic analysis of multi-focal (but monoclonal) cancers<sup>26</sup>, by FISH-based screening of tissue sections<sup>27–29</sup> or immuno-selected cells<sup>30</sup>, by the molecular probing of multiple small biopsies<sup>31</sup> or of micro-dissected tissue<sup>32–34</sup> and, recently, by sector-ploidy profiling<sup>35</sup>. Small numbers of individual circulating tumour cells have also been scrutinized for their divergent genetic profiles<sup>36,37</sup>. These studies collectively testify that contemporaneous intracolon genetic heterogeneity is commonplace and, in some cases at least, the degree of clonal diversity is predictive of disease progression<sup>31</sup>. Most of these data derive from epithelial carcinomas with complex genetic profiles, coupled, in most cases, to genetic instability. In such cases the historical timing and sequence of critical or driver mutational events is effectively buried and clonal architecture could be extremely complex unless clonal dominance occurs.

A common assumption for both leukaemias and cancer in general, based on the original evolutionary model of ref. 22, is that progression of disease and predominant genetic profiles reflect sequentially dominant clones and an essentially linear dynamic. Our data (summarized in Supplementary Fig. 1) suggest dynamic patterns of subclonal development and ancestral relationships that are nonlinear with a variable branching architecture. Patterns of genetic diversity in other cancers—assessed by single cell or ploidy sorted cell comparative genomic hybridization (CGH)<sup>35,38,39</sup>, oncogenically neutral microsatellite markers<sup>31,40</sup> or deep-sequenced IGH gene rearrangements<sup>41</sup>—also indicate nonlinear, branching clonal trajectories. Collectively, these data indicate that cancer has a cellular and genetic architecture reminiscent of Darwin’s iconic evolutionary tree (or bush) diagram depicting speciation<sup>42</sup>.

The extent of genetic variegation in subclones that we detect must be a significant underestimate. We screened for a limited number of pre-selected CNA, which means that other CNA plus any sequence-based driver mutations present will not have been registered. Moreover, antecedent or intermediary subclones, below the 1–2% frequency which we set as a threshold, were identified with more extensive screening in several cases (see Supplementary Fig. 2 patients 2, 5, 15, 18, 20, 6). Identifying the full complexity of subclonal architecture and genetic diversity in ALL (and other cancers) will ultimately require whole-genome analysis at the single cell level.

Our data provide the first direct evidence for genetic diversity of cancer propagating cells within individual patients. The consistent patterns of subclonal regeneration in mice (that is, in different mice injected with the same cellular inoculum) suggest variable capacity intrinsically associated with the distinct genotypes of propagating

cells. The competitive potency of particular subclones observed, however, may to some extent reflect selective pressures exerted by regenerative stress in a murine tissue environment. Natural clonal selection in patients might produce different outcomes.

It will be important to assess if genetic diversity of propagating cells holds true for other types of leukaemia and cancer in general. If it does, then there would be significant implications for both the cancer-stem-cell concept itself and for the therapeutic targeting of such cells. The original model of a distinct, hierarchically positioned subpopulation of cancer stem cells<sup>5</sup> has proved contentious in both ALL<sup>15</sup> and other cancers<sup>43–48</sup>. It has been suggested that the NOD/SCID *in vivo* readout for human cancer stem cell may, at least for some cancers, simply register dominant subclones<sup>43,48</sup>. Or, alternatively, that cancer stem cells exist but evolve over time<sup>44,45</sup>. We have previously documented that 'pre-leukaemic' and overt leukaemia propagating cells in ETV6–RUNX1-positive ALL, although clonally related by descent, are distinctive in IgH rearrangements and phenotype<sup>17</sup>. Our current data fit best with what we refer to as a 'back to Darwin' model for cancer propagating cells and resultant clonal architecture<sup>46</sup>. In this, cells with self-renewing properties have variegated genotypes providing the units of selection in the evolutionary diversification and progression of disease. Both sequential and concurrent genotypic variation in propagating cells occur in ALL and, we predict, are likely to do so in other cancers, providing a rich substrate for progression of disease. Although it has yet to be evaluated, it is likely that genetic diversity of cancer propagating cells will be associated with both frequency variation and diversity of functional properties, for example, differentiation status, niche occupancy, quiescence and drug or irradiation sensitivity. This may help to explain some of the inconsistencies and controversies in the cancer-stem-cell field<sup>44,47,48,49</sup>. Genetic diversity in cancer varies in extent with stage of disease<sup>28,33</sup>, probably reflecting the impact of intraclonal competition and ecological bottlenecks. Single cells might negotiate very stringent bottlenecks but the genetic profiles that we observed in relapsed ALL and as recorded in, for example, prostate cancer metastases<sup>37,50</sup> indicate continued diversification of propagating cells and dominant or therapy-resistant subclones.

This perspective contrasts with the unidimensional or flat (albeit very complex) genetic landscapes of cancer implied in portraits derived from whole-genome scans. This architectural distinction may be of some clinical consequence. Targeted therapy, if directed at mutant molecules, may have limited efficacy if the targets themselves are not initiating lesions but secondary mutations segregated in subclones, even when the latter appear dominant. Genetic variegation of cancer propagating cells may represent a significant roadblock to effective therapy.

## METHODS SUMMARY

Archival methanol:acetic-acid-fixed cytogenetic pellets from patients with ETV6–RUNX1 fusion-gene-positive ALL were obtained from several UK hospitals, with local ethical review committee approval (CCR 2285, Royal Marsden Hospital NHS Foundation Trust). The clinical and cytogenetic data on these patients are given in Supplementary Table 1. Interphase FISH was performed as previously described<sup>9,19</sup>.

In each case, at least 200 nuclei were scored for the presence of the ETV6–RUNX1 fusion gene in combination with hemizygous or homozygous deletion of ETV6, RUNX1, PAX5, CDKN2A and BTG1 and 11q, as well as gain of RUNX1 and duplication of the ETV6–RUNX1 fusion gene. Controls included the scoring of residual normal cells within the diagnostic sample and scoring leukaemic cells with probes hybridizing to irrelevant oncogenes (BCR, ABL) (see main Methods and Supplementary Figs 9 and 10).

NOD/SCID IL2R $\gamma$ <sup>null</sup> mice that lack any B, T and natural killer cell activity were bred and maintained under sterile conditions in accordance with Home Office regulations. Transplantation of cells was by intra-tibial injections in 7–14-week-old mice after 250 cGy irradiation. Peripheral engraftment was assessed at 9–10 weeks after transplantation and if >2% mice were killed. Further analysis included the assessment of bone marrow/spleen engraftment, FISH analysis, histological analysis and serial transplantation. For serial transplantations, recovered bone marrow cells were stained with human CD45 to detect human engraftment.

An equivalent of  $2 \times 10^3$  to  $2 \times 10^5$  human cells was transplanted by intra-tibial injections.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 20 May; accepted 27 October 2010.

Published online 15 December 2010.

- Greenman, C. *et al.* Patterns of somatic mutation in human cancer genomes. *Nature* **446**, 153–158 (2007).
- Beroukhi, R. *et al.* The landscape of somatic copy-number alteration across human cancers. *Nature* **463**, 899–905 (2010).
- Fox, E. J., Salk, J. J. & Loeb, L. A. Cancer genome sequencing—an interim analysis. *Cancer Res.* **69**, 4948–4950 (2009).
- Marusyk, A. & Polyak, K. Tumor heterogeneity: causes and consequences. *Biochim. Biophys. Acta* **1805**, 105–117 (2010).
- Dick, J. E. Stem cell concepts renew cancer research. *Blood* **112**, 4793–4807 (2008).
- Klein, C. A. *et al.* Genetic heterogeneity of single disseminated tumour cells in minimal residual cancer. *Lancet* **360**, 683–689 (2002).
- Greaves, M. F. & Wiemels, J. Origins of chromosome translocations in childhood leukaemia. *Nature Rev. Cancer* **3**, 639–649 (2003).
- Mullighan, C. G. *et al.* Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* **446**, 758–764 (2007).
- Bateman, C. M. *et al.* Acquisition of genome-wide copy number alterations in monozygotic twins with acute lymphoblastic leukemia. *Blood* **115**, 3553–3558 (2010).
- Loncarevic, I. F. *et al.* Trisomy 21 is a recurrent secondary aberration in childhood acute lymphoblastic leukemia with TEL/AML1 fusion. *Genes Chromosom. Cancer* **24**, 272–277 (1999).
- Kitagawa, Y. *et al.* Prevalent involvement of illegitimate V(D)J recombination in chromosome 9p21 deletions in lymphoid leukemia. *J. Biol. Chem.* **277**, 46289–46297 (2002).
- Mullighan, C. G. *et al.* BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros. *Nature* **453**, 110–114 (2008).
- Feldhahn, N. *et al.* Activation-induced cytidine deaminase acts as a mutator in BCR-ABL1-transformed acute lymphoblastic leukemia cells. *J. Exp. Med.* **204**, 1157–1166 (2007).
- van Dongen, J. J. M., Szczepanski, T. & Adriaansen, H. J. In *Leukemia* (eds Henderson, E. S., Lister, T. A. & Greaves, M. F.) 85–129 (Saunders, 2002).
- le Viseur, C. *et al.* In childhood acute lymphoblastic leukemia, blasts at different stages of immunophenotypic maturation have stem cell properties. *Cancer Cell* **14**, 47–58 (2008).
- Castor, A. *et al.* Distinct patterns of hematopoietic stem cell involvement in acute lymphoblastic leukemia. *Nature Med.* **11**, 630–637 (2005).
- Hong, D. *et al.* Initiating and cancer-propagating cells in TEL-AML1-associated childhood leukemia. *Science* **319**, 336–339 (2008).
- Breathnach, F., Chessells, J. M. & Greaves, M. F. The aplastic presentation of childhood leukaemia: a feature of common-ALL. *Br. J. Haematol.* **49**, 387–393 (1981).
- Horsley, S. W. *et al.* Genetic lesions in a preleukemic aplasia phase in a child with acute lymphoblastic leukemia. *Genes Chromosom. Cancer* **47**, 333–340 (2008).
- Zuna, J. *et al.* TEL deletion analysis supports a novel view of relapse in childhood acute lymphoblastic leukemia. *Clin. Cancer Res.* **10**, 5355–5360 (2004).
- Mullighan, C. G. *et al.* Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. *Science* **322**, 1377–1380 (2008).
- Nowell, P. C. The clonal evolution of tumor cell populations. *Science* **194**, 23–28 (1976).
- Gatenby, R. A. & Vincent, T. L. An evolutionary model of carcinogenesis. *Cancer Res.* **63**, 6212–6220 (2003).
- Merlo, L. M. F., Pepper, J. W., Reid, B. J. & Maley, C. C. Cancer as an evolutionary and ecological process. *Nature Rev. Cancer* **6**, 924–935 (2006).
- Teixeira, M. R. *et al.* Karyotypic comparisons of multiple tumorous and macroscopically normal surrounding tissue samples from patients with breast cancer. *Cancer Res.* **56**, 855–859 (1996).
- Takahashi, T. *et al.* Clonal and chronological genetic analysis of multifocal cancers of the bladder and upper urinary tract. *Cancer Res.* **58**, 5835–5841 (1998).
- Cottu, P. H. *et al.* Intratumoral heterogeneity of HER2/neu expression and its consequences for the management of advanced breast cancer. *Ann. Oncol.* **19**, 596–597 (2008).
- Park, S. Y. *et al.* Cellular and genetic diversity in the progression of in situ human breast carcinomas to an invasive phenotype. *J. Clin. Invest.* **120**, 636–644 (2010).
- Clark, J. *et al.* Complex patterns of ETS gene alteration arise during cancer development in the human prostate. *Oncogene* **27**, 1993–2003 (2008).
- Shipitsin, M. *et al.* Molecular definition of breast tumor heterogeneity. *Cancer Cell* **11**, 259–273 (2007).
- Maley, C. C. *et al.* Genetic clonal diversity predicts progression to esophageal adenocarcinoma. *Nature Genet.* **38**, 468–473 (2006).
- Aubele, M. *et al.* Intratumoral heterogeneity in breast carcinoma revealed by laser-microdissection and comparative genomic hybridization. *Cancer Genet. Cytogenet.* **110**, 94–102 (1999).
- Boland, C. R. *et al.* Microallelotyping defines the sequence and tempo of allelic losses at tumour suppressor gene loci during colorectal cancer progression. *Nature Med.* **1**, 902–909 (1995).

34. Geyer, F. C. *et al.* Molecular analysis reveals a genetic basis for the phenotypic diversity of metaplastic breast carcinomas. *J. Pathol.* **220**, 562–573 (2010).
35. Navin, N. *et al.* Inferring tumor progression from genomic heterogeneity. *Genome Res.* **20**, 68–80 (2010).
36. Stoecklein, N. H. *et al.* Direct genetic analysis of single disseminated cancer cells for prediction of outcome and therapy selection in esophageal cancer. *Cancer Cell* **13**, 441–453 (2008).
37. Attard, G. *et al.* Characterization of *ERG*, *AR* and *PTEN* gene status in circulating tumor cells from patients with castration-resistant prostate cancer. *Cancer Res.* **69**, 2912–2918 (2009).
38. Klein, C. A. & Stoecklein, N. H. Lessons from an aggressive cancer: evolutionary dynamics in esophageal carcinoma. *Cancer Res.* **69**, 5285–5288 (2009).
39. Kuukasjärvi, T. *et al.* Genetic heterogeneity and clonal evolution underlying development of asynchronous metastasis in human breast cancer. *Cancer Res.* **57**, 1597–1604 (1997).
40. Tsao, J.-L. *et al.* Colorectal adenoma and cancer divergence. Evidence of multilineage progression. *Am. J. Pathol.* **154**, 1815–1824 (1999).
41. Campbell, P. J. *et al.* Subclonal phylogenetic structures in cancer revealed by ultra-deep sequencing. *Proc. Natl Acad. Sci. USA* **105**, 13081–13086 (2008).
42. Barrett, P. H. *et al.* (eds.) *Charles Darwin's Notebooks, 1836–1844* (Cambridge Univ. Press, 1987).
43. Adams, J. M. & Strasser, A. Is tumor growth sustained by rare cancer stem cells or dominant clones? *Cancer Res.* **68**, 4018–4021 (2008).
44. Visvader, J. E. & Lindeman, G. J. Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. *Nature Rev. Cancer* **8**, 755–768 (2008).
45. Rosen, J. M. & Jordan, C. T. The increasing complexity of the cancer stem cell paradigm. *Science* **324**, 1670–1673 (2009).
46. Greaves, M. Cancer stem cells: back to Darwin? *Semin. Cancer Biol.* **20**, 65–70 (2010).
47. Maenhaut, C., Dumont, J. E., Roger, P. P. & van Staveren, W. C. G. Cancer stem cells: a reality, a myth, a fuzzy concept or a misnomer? An analysis. *Carcinogenesis* **31**, 149–158 (2010).
48. Shackleton, M., Quintana, E., Fearon, E. R. & Morrison, S. J. Heterogeneity in cancer: cancer stem cells versus clonal evolution. *Cell* **138**, 822–829 (2009).
49. Polyak, K. Breast cancer: origins and evolution. *J. Clin. Invest.* **117**, 3155–3163 (2007).
50. Liu, W. *et al.* Copy number analysis indicates monoclonal origin of lethal metastatic prostate cancer. *Nature Med.* **15**, 559–565 (2009).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This work is supported by specialist programme grants from The Kay Kendall Leukaemia Fund (M.G.) and Leukaemia & Lymphoma Research (M.G., T.E.) and a Deutsche Forschungsgemeinschaft fellowship LU 1474/1-1 (to C.L.). T.E. and C.L. acknowledge support from the Oxford BRC.

**Author Contributions** K.A. carried out the FISH analyses. C.L. and Y.G. conducted the *in vivo* experiments. C.M.B. analysed the SNP array and FISH analysis of the patient with aplasia and ALL. S.M.C. and I.T. performed cell immunostaining and sorting. F.W.v.D. provided SNP array data. H.K., A.V.M. and J.S. provided patient data and samples. T.E. advised on design and interpretation of *in vivo* experiments. L.K. supervised the FISH studies. L.K. and M.G. designed the overall study. M.G. wrote the paper with critical input from T.E., L.K. and other authors.

**Author Information** SNP array data have been deposited in the Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/>) under accession code GSE24412. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to M.G. ([mel.greaves@icr.ac.uk](mailto:mel.greaves@icr.ac.uk)).



## METHODS

**Interphase fluorescence *in situ* hybridization (FISH).** Archival methanol:acetic acid-fixed cytogenetic pellets from patients with *ETV6-RUNX1* fusion-gene-positive ALL were obtained from several UK hospitals, with local ethical review committee approval (CCR 2285, Royal Marsden Hospital NHS Foundation Trust). Interphase FISH for the *ETV6-RUNX1* fusion gene was performed using a commercial LSI *TEL-AML1* extra signal (ES) probe (Vysis, Abbott Laboratories Ltd) according to the manufacturers' instructions. This probe set contains a 350-kb probe for the 5' end of *ETV6* (exons 1–4) and a 500-kb probe covering the entire *RUNX1* gene. The FISH signal pattern for the *ETV6-RUNX1* fusion-gene-positive cells using the Vysis probe is two red (one large, one small *RUNX1* signals), one green (*ETV6* allele not involved in the translocation), one red/green (yellow) fusion signal corresponding to the *ETV6-RUNX1* fusion gene. Bacterial artificial chromosome (BAC) or fosmid probes for the *PAX5*, *CDKN2A*, *BTG1*, *TBL1XR1* genes, 11q and other regions of interest were obtained from the BACPAC Resource Centre, Children's Hospital, Oakland Research Institute (<http://bacpac.chori.org>). These were labelled by nick translation with biotin-16-dUTP or digoxigenin-11-dUTP (Roche) and hybridized in combination with the *ETV6-RUNX1* ES probe. FISH was performed by standard protocols<sup>9,19</sup> and labelled probes detected with streptavidin-Cy5 (biotinylated probes) and (1) monoclonal anti-digoxigenin (Sigma), (2) horse anti-mouse IgG-Texas red (Vector Laboratories) and (3) goat anti-horse IgG-Texas Red (Jackson Immunochemicals) (for digoxigenin-labelled probes). Fluorescent signals were viewed using an Olympus AX2 fluorescence microscope equipped with narrow bandpass filters for DAPI, FITC, Spectrum orange, Texas red and Cy5. Images were captured and analysed using a charge-coupled device (Photometrics) and SmartCapture 3 software version 3.0.4 (Digital Scientific).

**Establishing cutoff levels.** In each case, at least 200 nuclei were scored for the presence of the *ETV6-RUNX1* fusion gene in combination with hemizygous or homozygous deletion of *ETV6*, *RUNX1*, *PAX5*, *CDKN2A* and *BTG1* and 11q, as well as gain of *RUNX1* and the *ETV6-RUNX1* fusion gene. Diagnostic slides from 26 cases were assessed for hybridization efficiency by scoring the residual normal (*ETV6-RUNX1* fusion negative) cells on the same slide (see Supplementary Fig. 9). The percentage of these cells with loss of a single *CDKN2A* or *PAX5* signal was 0–3% (mean = 1%). As a further control, we hybridized a subset ( $n = 11$ ) of *ETV6-RUNX1* fusion-gene-positive cases with uninvolved oncogene probes *BCR* and *ABL* (see Supplementary Fig. 10). The percentage of cells with the expected normal signal pattern (two red, two green) was 96–99% (mean = 98%). Cutoff levels for each probe were established by three-colour FISH (test probe in combination with the *ETV6-RUNX1* ES probe) using three normal control peripheral blood slides per probe. As *ETV6* and *RUNX1* were scored on each slide, the values for these two probes were based on 12 slides in total. We used a cutoff = mean + 2 × standard deviation. Cutoff levels for each probe (in combination with the *ETV6-RUNX1* fusion) were established using 3–12 normal control peripheral blood slides. The cutoff levels for three-colour FISH experiments are given in Supplementary Table 2. Cutoff levels for four-colour FISH were the same as above, except for Texas red probes. The cutoff for loss of one signal using a fourth probe detected with Texas red was 6.9%, because of spectral overlap between Texas red (used to detect digoxigenin-labelled probes) and Spectrum orange (used to label *RUNX1* in the commercial *ETV6-RUNX1* ES probe). However, in most cases we used three-colour FISH as a cross-check to infer whether clones below this cutoff were real (Supplementary Fig. 2). Only fusion-gene-positive cells that also showed the small extra red signal (generated by disruption of *RUNX1*) were used to calculate the relative frequencies of the various subclones.

**Genome mapping analysis.** Mapping analysis was performed using 500 ng of tumour DNA. DNA was prepared according to manufacturer's instructions using the GeneChip mapping 500K assay protocol for hybridization to GeneChip Mapping 250K Nsp and Sty arrays (Affymetrix). Briefly, genomic DNA was digested in parallel with restriction endonucleases NspI and StyI, ligated to an adaptor, and subjected to polymerase chain reaction (PCR) amplification with adaptor-specific primers. The PCR products were digested with DNaseI and labelled with a biotinylated nucleotide analogue. The labelled DNA fragments were hybridized to the microarray, stained by streptavidin-phycoerythrin

conjugates, and washed using the Affymetrix Fluidics Station 450 then scanned with a GeneChip scanner 3000 7G.

**Copy number and LOH analysis.** SNP genotypes were obtained using Affymetrix GCOS software (version 1.4) to obtain raw feature intensity and Affymetrix GTYPE software (version 4.0) using the BRLMM algorithm to derive SNP genotypes. The samples were analysed using CNAG 3.0 (<http://plaza.umin.ac.jp/genome>), comparing tumour sample with unpaired control DNA to determine copy number and loss of heterozygosity (LOH) caused by imbalance<sup>51</sup>. The position of regions of LOH were identified using the University of California Santa Cruz (UCSC) Genome Browser, May 2004 Assembly (<http://genome.ucsc.edu/cgi-bin/hgGateway>).

**Combined fluorescence immunophenotype and FISH.** Bone marrow and spleen cells from leukaemic mice were cytospun and used for combined or triple-colour immunophenotype/FISH analysis as previously described<sup>52</sup>. Briefly, cells were air dried and fixed in acetone before incubating in primary biotinylated mouse anti-human CD45 (Clone F10-89-4) and detecting with Avidin-AMCA (Vector Laboratories). After antibody staining, the slides were hybridized with the Vysis *ETV6-RUNX1* ES fusion gene FISH probe as described above. Cells were viewed using a Zeiss Axioskop fluorescence microscope fitted with a dual bandpass FITC and rhodamine filter, as well as individual DAPI (for AMCA immunophenotype), FITC, rhodamine and Cy5. Images were captured using a charge-coupled device (Photometrics) and fluorescence signals merged and analysed using SmartCapture X software version 2.6.2 (Digital Scientific).

**Cell separation, phenotyping and sorting.** Total mononuclear cells were isolated by Ficoll gradient centrifugation and directly cryopreserved in DMSO for later use. After thawing dead cells were evaluated and excluded by FACS after staining with Hoechst 33258 (Invitrogen). For sorting CD34 and CD20 positive and negative subsets, samples were stained with either mouse anti-human CD34 (IgG<sub>1</sub>, Dako) or CD20 (IgG<sub>1</sub>, Southern Biotech) followed by anti-mouse IgG labelled with Pacific blue (Invitrogen). Nonspecific binding of antibody was assessed using mouse IgG<sub>1</sub> isotype control stained samples (Dako). Sorting was performed on a BD FACSAria with analysis by Becton Dickinson FACSDiVa software. Samples were gated on forward- and side-scatter plots for mononuclear cells and further gated in forward-scatter height versus area to exclude clumped cells. Before xeno-transplantation, some cells were stained with anti-CD19 PE (BD Pharmingen), CD34 FITC (BD Pharmingen) and CD38 APC (BD Pharmingen). CD34<sup>+</sup>CD38<sup>low</sup>CD19<sup>+</sup> and pro-B CD34<sup>+</sup>CD38<sup>+</sup>CD19<sup>+</sup> cells were purified by flow cytometry (in this case, using MoFlo, Dako). Data acquisition and analysis were done with Summit (Dako) software. For multi-colour cell sorting 'fluorescence minus one' controls were used to determine positive and negative staining boundaries<sup>53</sup>. Human cells regenerating in mice were identified by staining with anti-CD45 PeCy7 (BD Pharmingen).

**NOD/SCID mouse transplantation.** NOD/SCID IL2Rγ<sup>null</sup> mice that lack any B, T and natural killer cell activity were bred and maintained at the Weatherall Institute of Molecular Medicine animal facility in accordance with Home Office regulations. Animals were handled under sterile conditions. Transplantations of  $2 \times 10^3$ – $10^6$  cells were performed by intra-tibial injections in 7–14-week-old mice. Recipients received 250 cGy of total body irradiation before cell injection. Peripheral engraftment was assessed at 9–10 weeks after transplantation and if peripheral engraftment was >2% mice were killed. Further analysis included the assessment of bone marrow/spleen engraftment, FISH analysis, histological analysis and serial transplantation. For serial transplantations recovered bone marrow cells were stained with human CD45 to detect human engraftment. An equivalent of  $2 \times 10^3$  to  $2 \times 10^5$  human cells was transplanted by intra-tibial injections.

**May-Grünwald Giemsa staining.** The histological analysis of patient samples and mouse bone marrow was performed by May-Grünwald Giemsa staining of bone marrow smears, bone marrow cytospin preparations and spleen swabs. Slides were analysed on an Olympus BX60 microscope.

51. Nannya, Y. *et al.* A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. *Cancer Res.* **65**, 6071–6079 (2005).
52. Kearney, L. & Colman, S. in *Methods in Molecular Biology. Leukemia. Methods and Protocols* Vol. 538 (ed. So, C. W. E.) 57–70 (Humana Press, 2009).
53. Maecker, H. T. & Trotter, J. Flow cytometry controls, instrument setup, and the determination of positivity. *Cytometry A* **69**, 1037–1042 (2006).

# Atom-by-atom spectroscopy at graphene edge

Kazu Suenaga<sup>1</sup> & Masanori Koshino<sup>1</sup>

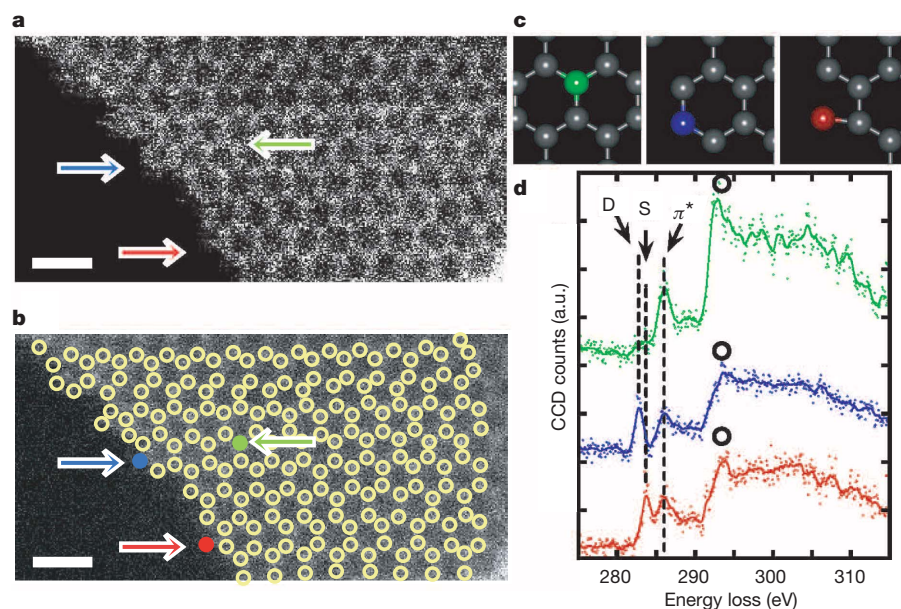
The properties of many nanoscale devices are sensitive to local atomic configurations, and so elemental identification and electronic state analysis at the scale of individual atoms is becoming increasingly important. For example, graphene is regarded as a promising candidate for future devices, and the electronic properties of nanodevices constructed from this material are in large part governed by the edge structures<sup>1</sup>. The atomic configurations at graphene boundaries have been investigated by transmission electron microscopy and scanning tunnelling microscopy<sup>2–4</sup>, but the electronic properties of these edge states have not yet been determined with atomic resolution. Whereas simple elemental analysis at the level of single atoms can now be achieved by means of annular dark field imaging<sup>5</sup> or electron energy-loss spectroscopy<sup>6,7</sup>, obtaining fine-structure spectroscopic information about individual light atoms such as those of carbon has been hampered by a combination of extremely weak signals and specimen damage by the electron beam. Here we overcome these difficulties to demonstrate site-specific single-atom spectroscopy at a graphene boundary, enabling direct investigation of the electronic and bonding structures of the edge atoms—in particular, discrimination of single-, double- and triple-coordinated carbon atoms is achieved with atomic resolution. By demonstrating how rich chemical information can be obtained from single atoms through energy-loss near-edge fine-structure analysis<sup>8</sup>, our results should open the way to exploring the local electronic structures of various nanodevices and individual molecules.

A low-voltage scanning transmission electron microscope (STEM) was used for the single-atom spectroscopy<sup>9</sup>. Flakes were cleaved from the synthetic highly oriented pyrolytic graphite (HOPG) and put onto the microgrids for energy-loss near-edge fine structure (ELNES)

analysis. STEM annular dark field (ADF) images indicate that the graphene flakes have open and active edges<sup>3</sup> and that the edges are steadily etched by the incident electron beam when the probe-scanning is repeated at the same region (Supplementary Fig. 1). The accelerating voltage used here (60 kV) is below the critical energy predicted for severe knock-on damage<sup>10</sup> and therefore the carbon atoms in bulk are mostly stable. Only the edge atoms are mobile during the observation, as indicated by the wiggling contrast frequently observed at the edge regions. The fast Fourier transformation of an ADF image of few-layer graphene shows that the spatial resolution of the experimental set-up is better than 0.106 nm (inset to Supplementary Fig. 1a) and so the hexagonal network of carbon atoms, separated by about 0.14 nm, is clearly visible in a monolayer region (Supplementary Fig. 1b). A probe of the same size and brightness was used for the following ELNES analysis.

Figure 1a shows a typical ADF image of the edge region of a single graphene layer. The hexagonal network of carbon atoms in bulk is visible on the right-hand side of the image and the vacuum region appears in black on the left-hand side. The possible carbon atom positions derived from the local intensity maxima of ADF signals are marked by yellow circles after an image-smoothing process in Fig. 1b. There is strong wiggle contrast at the edge regions and some of the atom positions cannot be completely identified. We note that some of the hexagonal networks are imperfect and considerably reconstructed at the edge region.

The typical ELNES spectra of carbon K (1s)-edge are displayed with their corresponding atomic positions in Fig. 1c. Figure 1d shows three characteristic carbon K-edge fine structures extracted using sequential electron energy-loss spectroscopy (EELS) with probe-scanning (known as the spectrum-image mode)<sup>11</sup>. The spectrum in green was recorded at an atomic position in bulk (indicated by a green circle and arrow in



**Figure 1 | Graphene edge spectroscopy.** **a**, ADF image of single graphene layer at the edge region. No image-processing has been done. Atomic positions are marked by circles in a smoothed image (**b**). Scale bars, 0.5 nm. **c**, ELNES of carbon K (1s) spectra taken at the colour-coded atoms indicated in **b**. Green, blue and red spectra correspond to the normal  $sp^2$  carbon atom, a double-coordinated atom and a single-coordinated atom, respectively. These different states of atomic coordination are marked by coloured arrows in **a** and **b** and illustrated in **c**. CCD, charge-coupled device.

<sup>1</sup>Nanotube Research Center, National Institute of Advanced Industrial Science and Technology (AIST), AIST Central 5, Tsukuba 305-8565, Japan.



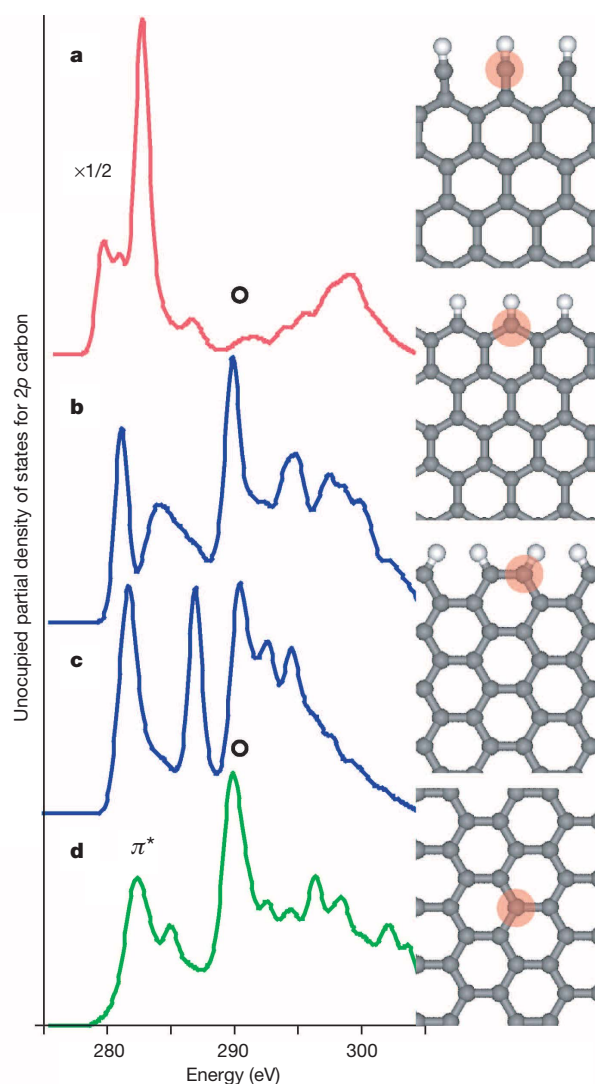
Fig. 1b) as a reference. This spectrum exhibits the features of typical  $sp^2$  coordinated carbon atoms, such as the sharp  $\pi^*$  peak around 286 eV and the exciton peak of  $\sigma^*$  at 292 eV. These features are in good agreement with the previously reported spectra recorded from a bulk graphite specimen<sup>12</sup>. The spectrum in blue was recorded from an edge atom located at the border of the hexagonal network with two-coordination, as illustrated in Fig. 1c. Remarkably, this spectrum has an extra peak around  $282.6 \pm 0.2$  eV (labelled D in Fig. 1d), with the  $\pi^*$  peak having reduced intensity. Also the exciton peak intensity is considerably reduced and broadened compared to the bulk spectrum (marked by open circles).

The spectrum in red shows similar features, also with weaker  $\pi^*$  peak and broadened  $\sigma^*$  peak. Its extra peak occurs at a different energy position of  $283.6 \pm 0.2$  eV (labelled S in Fig. 1d). It is extremely difficult to assign the atomic position completely for this red spectrum because the spectrum disappears quickly and is not fully reproducible. The edge region of the specimen tends to be strongly damaged and the edge morphology frequently changes after recording the spectrum image. Therefore, we can reasonably infer that this energy state must be somehow damage-related. One of the possible models for this edge structure is the Klein edge<sup>13,14</sup>. The edge atom indicated in red in Fig. 1b is indeed single-bonded to its neighbour. The structure should be very unstable under the incident electron beam and so it may also explain the wiggling contrast often observed at the graphene edge.

These spectral features involving peaks D and S have not previously been reported, to our knowledge. No fingerprinting method, comparing against the reference spectra of the existing polymorphic carbon, is able to explain them. We therefore performed ELNES simulations to correlate the experimental features with different atomic configurations (Fig. 2). The  $\pi^*$  peak shift to the lower energy is well reproduced for the edge atoms in the Klein, zigzag and armchair edge configurations (Fig. 2a, b and c), in comparison with the bulk carbon atom (Fig. 2d). The diminished excitonic effect can be confirmed for the Klein edge (Fig. 2a). The peak shift around 2 eV is well reproduced for the zigzag edge (Fig. 2b). In the spectrum of the armchair edge a sharp peak between  $\pi^*$  and  $\sigma^*$  is expected (Fig. 2c).

To show an atom-by-atom spectroscopy, we also performed EELS in the spectrum-line mode across a graphene edge. The probe scanned across the protruded carbon atom—the Klein edge—from the vacuum to the bulk region along the dotted line in Fig. 3a. A series of 100 spectra were sequentially recorded by scanning the electron probe with a constant step of about 0.02 nm. The total acquisition time was as small as 50 s. The illustrated model in Fig. 3b shows that eight carbon atoms were investigated in the spectrum line. Figure 3c shows a profile of ADF signals (in red) that was simultaneously recorded with the ELNES spectra. It shows good agreement with the simulated profile (in blue) showing eight maxima sequentially corresponding to the eight carbon atoms. Although the experimental profile is rather scattered owing to specimen instability or a possible inclination of the specimen to the incident electron beam, which should produce a slight asymmetry in the profile of the carbon doublets, we can deduce the carbon atomic positions reasonably well from the line profile and extract the ELNES spectra corresponding to each atom. Figure 3d shows ELNES fine structures obtained in this way, with the corresponding atoms numbered in Fig. 3c (each spectrum presented consists of four spectra in total).

The delocalization effect at the carbon  $K$ -edge ( $\sim 290$  eV) with an incident electron probe of 60 kV is estimated as 0.20–0.25 nm in classical theory<sup>8</sup> and as  $\sim 0.12$  nm at 300 kV more recently<sup>15</sup>. Therefore the EELS signals, if combined with the probe size ( $\sim 0.1$  nm), may not be completely localized at the single atoms on which the probe is exactly positioned. However the series of ELNES spectra in Fig. 3d strongly suggest that site-specific spectroscopy is indeed possible with atomic resolution at the graphene edge. The spectrum from atom 1 clearly shows peak S at 283.6 eV (indicated by a dotted circle), which is related to the Klein edge, but spectrum 2 does not show peak S (it may be a minor feature). Spectrum 5 shows a small trace of peak D, which can reasonably be explained by a possible introduction of the bond-breakage



**Figure 2 | ELNES simulations for three graphene edge structures.** Carbon  $K$ -edge spectra simulated for the Klein edge (a), zigzag edge (b), armchair edge (c) and bulk (three-coordinated) atom (d). A core-hole was introduced by partially removing a  $1s$  electron from the carbon atoms (indicated by pink shading) to estimate the relative peak shift of the spectra. The reduced exciton peak found experimentally is well reproduced. The simulated ELNES from the zigzag and armchair edges show at least a qualitative match with experiments, although the absolute value for the energy shift cannot be fully confirmed.

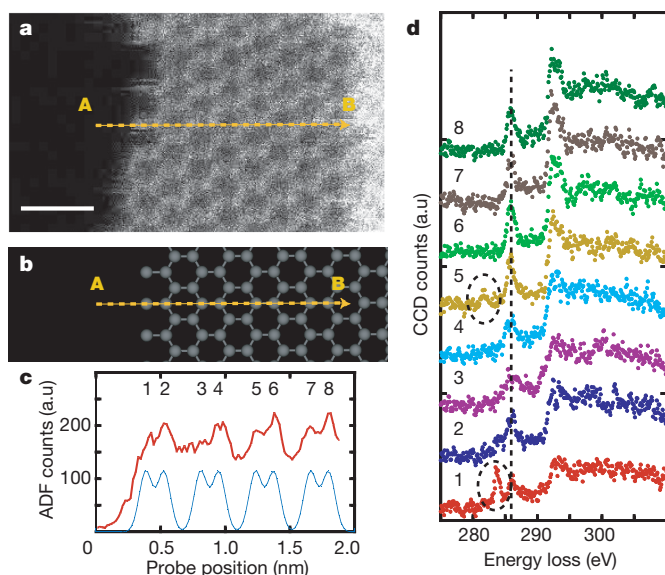
during the probe-scanning across the atom. Spectrum 8 from an atom 1.5 nm away from the edge shows normal  $sp^2$  features with the sharp  $\pi^*$  and excitonic  $\sigma^*$  peaks, which is very close to the bulk spectrum<sup>12</sup>.

We performed intensity mapping of peaks D and S to estimate the delocalization effects further. A number of experiments, involving one set of spectrum-image and seven sets of spectrum-line on the graphene edges, are summarized in Supplementary Figs 3, 4 and 5. Results confirm that single-atom spectroscopy at specific sites of the graphene edge is indeed feasible with the reduced delocalization effect.

We found no trace of oxygen at the investigated edges. This may contradict a generally accepted concept in which the graphene edge can be terminated by -OH or -COOH groups and the edge carbon atoms cannot be bared<sup>1</sup>. In this experiment, *in situ* etching with continuous removal of the carbon edge atoms in vacuum always takes place and therefore the edge structures are always kept fresh.

From this study, we have picked up some practical information about the graphene edge engineering. The open edges involve both single- and double-coordinated carbon atoms but their specific edge





**Figure 3 | Atom-by-atom spectroscopy across the Klein edge.** **a**, ADF image of graphene edge (no image-processing). The dotted arrow indicates where the spectrum-line was made (A to B). Scale bar, 0.5 nm. **b**, An atomic model of the investigated edge. **c**, Line-profile of the ADF counts (in red) recorded simultaneously with the spectrum-line. For comparison with the simulated ADF counts (blue), the number of each atom is indicated (from 1 to 8). **d**, The carbon K-edge ELNES obtained from each atom across the Klein edge. The single-coordinated carbon atom (numbered 1) clearly shows peak S.

states are completely localized at the atomic level. Even for triple-coordinated carbon atoms, slight electronic structure modification, as indicated by the restricted excitonic effect (or the reduced  $\sigma^*$  peak), may exist near the edge region but it vanishes after 1.5 nm from the edge front. The properties of graphene nanoribbons with smaller widths might be governed by the edge effects<sup>16</sup>.

It is very surprising that the EELS signal delocalization has turned out not to be very important for atom-by-atom spectroscopy in the present experiment. The EELS signal delocalization should be substantially decreased when a lower accelerating voltage is used for the incident electron probe<sup>8</sup>. The delocalization effect with a 30–60 kV incident probe is only a fraction of that for the normal STEM operation voltage at 200–300 kV. Lowering the accelerating voltage of the electron microscope is therefore very beneficial, reducing the delocalization effect in addition to contrast enhancement and damage reduction.

ELNES analysis from single atoms is highly desirable because the rich information it supplies will become accessible from individual atoms at any local area. The ELNES fingerprinting method has been widely used to determine the electronic/bonding states of unknown materials by comparison with the reference spectra of known materials. For example, the chemical state of  $\text{Ce}^{3+}$  or  $\text{Ce}^{4+}$  in metallofullerene molecules has been clearly discriminated at the single-atom level simply by measuring the energy shift<sup>17</sup>. Here we have demonstrated the possibilities of ELNES spectra analysis beyond the simple fingerprinting method. Non-bulk atoms provide peculiar electronic structures and therefore their ELNES should be completely new (or previously unknown) and cannot be compared with any existing reference. Further efforts should be made to obtain the electronic state information from new ELNES spectra by combining atomic resolution imaging with theoretical calculations.

## METHODS SUMMARY

**STEM-EELS experiments.** A JEOL 2100F transmission electron microscope with the DELTA corrector was operated at 60 kV (ref. 9). The energy resolution was around 0.4 eV. We used a probe of 0.1 nm diameter with 20 pA for experiments. For spectroscopy, we used GIF Quantum<sup>18</sup>, designed for low-voltage operations. The convergence angle for incident probe was set to 30 mrad, while the inner angle for ADF imaging was around 45–50 mrad, which is equal to the EELS collection angle. ELNES analysis was performed at each pixel while the incident probe

digitally scanned<sup>11</sup>. The spectrum-image mode, consisting of a two-dimensional set of ELNES spectra, takes longer for total acquisition and easily leads to the destruction of the specimen. Therefore we frequently used the spectrum-line mode, consisting of a one-dimensional set of ELNES spectra, in this study. Typical acquisition time is around 0.1 to 1.0 s for each spectrum. A spectrum line consists of 100 spectra, while an image spectrum consists of typically  $12 \times 12$  spectra (see also Supplementary Fig. 3).

**Specimen preparation.** Commercially available synthetic HOPG (NT-MDT Company) was used for experiments. Some of the flakes were cleaved using Scotch tapes and then transferred to transmission electron microscope microgrids following the method developed by Meyer and co-workers<sup>19</sup>.

**ELNES simulations.** The first-principles calculation based on DFT theory was used to estimate energy levels and partial density of states on carbon atoms of graphene structures. In the discrete variance- $X\alpha$  method, the energy levels and partial density of states of unoccupied carbon 2p orbitals are estimated from the self-consistent charge calculation. To estimate the threshold energy of the carbon K-edge, the core-hole effect was considered by employing the transition-state approximation method, which configures a half-electron removed from the carbon 1s orbital and added to an unoccupied orbital<sup>20,21</sup>. See also Supplementary Fig. 2.

Received 23 July; accepted 9 November 2010.

Published online 15 December 2010.

- Kobayashi, Y., Fukui, K., Enoki, T. & Kusakabe, K. Edge state on hydrogen-terminated graphite edges investigated by scanning tunneling microscopy. *Phys. Rev. B* **73**, 125415 (2006).
- Gass, M. H. et al. Free-standing graphene at atomic resolution. *Nature Nanotechnol.* **3**, 676–681 (2008).
- Liu, Z., Suenaga, K., Harris, P. & Iijima, S. Open and closed edges of graphene layers. *Phys. Rev. Lett.* **102**, 015501 (2009).
- Girit, C. O. et al. Graphene at the edge: stability and dynamics. *Science* **323**, 1705–1708 (2009).
- Krivanek, O. L. et al. Atom-by-atom structural and chemical analysis by annular dark-field electron microscopy. *Nature* **464**, 571–574 (2010).
- Suenaga, K. et al. Element selective single atom imaging. *Science* **290**, 2280–2282 (2000).
- Krivanek, O. L. et al. Gentle STEM: ADF imaging and EELS at low primary energies. *Ultramicroscopy* **110**, 935–945 (2010).
- Egerton, R. F. *Electron Energy-Loss Spectroscopy in the Electron Microscope* 2nd edn, 363–369 (Plenum, 1996).
- Sasaki, T. et al. Performance of low-voltage STEM/TEM with delta corrector and cold field emission gun. *J. Electron Microsc.* **59**, s7–s13 (2010).
- Zobelli, A., Gloter, A., Ewels, C. P., Seifert, G. & Colliex, C. Electron knock-on cross section of carbon and boron nitride nanotubes. *Phys. Rev. B* **75**, 245402 (2007).
- Jeanguillaume, C. & Colliex, C. Spectrum-image: the next step in EELS digital acquisition and processing. *Ultramicroscopy* **28**, 252–257 (1989).
- Garvie, L. A., Craven, A. J. & Brydson, R. Use of electron-energy loss near-edge fine structure in the study of minerals. *Am. Mineral.* **79**, 411–425 (1994).
- Klein, D. J. Graphitic polymer strips with edge states. *Chem. Phys. Lett.* **217**, 261–265 (1994).
- Kusakabe, K. & Maruyama, M. Magnetic nanographite. *Phys. Rev. B* **67**, 092406 (2003).
- Cosgriff, E. C., Oxley, M. P., Allen, L. J. & Pennycook, S. J. The spatial resolution of imaging using core-loss spectroscopy in the scanning transmission electron microscope. *Ultramicroscopy* **102**, 317–326 (2005).
- Jiao, L. et al. Facile synthesis of high-quality graphene nanoribbons. *Nature Nanotechnol.* **5**, 321–325 (2010).
- Suenaga, K. et al. Visualising and identifying single atoms using electron energy-loss spectroscopy with low accelerating voltage. *Nature Chem.* **1**, 415–418 (2009).
- Gubbens, A. et al. The GIF Quantum, a next generation post-column imaging energy filter. *Ultramicroscopy* **110**, 962–970 (2010).
- Pacile, D., Meyer, J. C., Girit, C. O. & Zettl, A. The two-dimensional phase of boron nitride: few-atomic-layer sheets and suspended membranes. *Appl. Phys. Lett.* **92**, 133107 (2008).
- Mizoguchi, T. et al. Core-hole effects on theoretical ELNES/NEXAFS of MgO. *Phys. Rev. B* **61**, 2180–2187 (2000).
- Koshino, M., Kurata, H. & Isoda, S. Study of structures at the boundary and defects in organic thin films of perchlorocoronene by high-resolution and analytical transmission electron microscopy. *Ultramicroscopy* **110**, 1465–1474 (2010).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This work is partially supported by the JST-CREST programme. We thank C. Jin for discussions and H. Kobayashi for specimen preparations.

**Author Contributions** K.S. designed and performed experiments. K.S. and M.K. analysed data. M.K. performed simulations. K.S. and M.K. co-wrote the paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to K.S. ([suenaga-kazu@aist.go.jp](mailto:suenaga-kazu@aist.go.jp)).

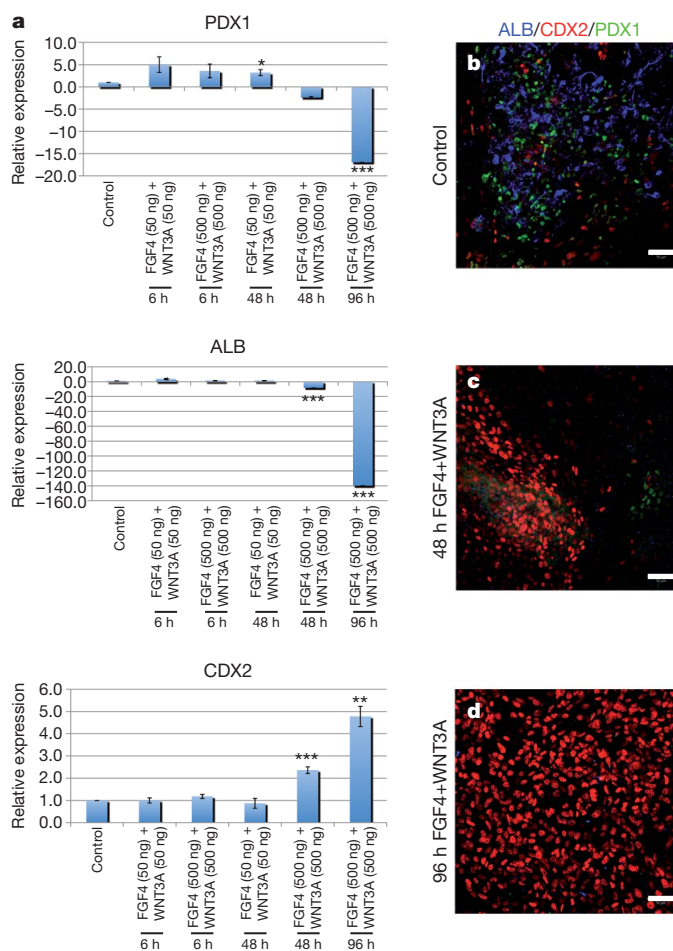
# Directed differentiation of human pluripotent stem cells into intestinal tissue *in vitro*

Jason R. Spence<sup>1</sup>, Christopher N. Mayhew<sup>1</sup>, Scott A. Rankin<sup>1</sup>, Matthew F. Kuhar<sup>1</sup>, Jefferson E. Vallance<sup>2</sup>, Kathryn Tolle<sup>1</sup>, Elizabeth E. Hoskins<sup>3</sup>, Vladimir V. Kalinichenko<sup>1,4</sup>, Susanne I. Wells<sup>3</sup>, Aaron M. Zorn<sup>1</sup>, Noah F. Shroyer<sup>1,2</sup> & James M. Wells<sup>1</sup>

Studies in embryonic development have guided successful efforts to direct the differentiation of human embryonic and induced pluripotent stem cells (PSCs) into specific organ cell types *in vitro*<sup>1,2</sup>. For example, human PSCs have been differentiated into monolayer cultures of liver hepatocytes and pancreatic endocrine cells<sup>3–6</sup> that have therapeutic efficacy in animal models of liver disease<sup>7,8</sup> and diabetes<sup>9</sup>, respectively. However, the generation of complex three-dimensional organ tissues *in vitro* remains a major challenge for translational studies. Here we establish a robust and efficient process to direct the differentiation of human PSCs into intestinal tissue *in vitro* using a temporal series of growth factor manipulations to mimic embryonic intestinal development<sup>10</sup>. This involved activin-induced definitive endoderm formation<sup>11</sup>, FGF/Wnt-induced posterior endoderm patterning, hindgut specification and morphogenesis<sup>12–14</sup>, and a pro-intestinal culture system<sup>15,16</sup> to promote intestinal growth, morphogenesis and cytodifferentiation. The resulting three-dimensional intestinal ‘organoids’ consisted of a polarized, columnar epithelium that was patterned into villus-like structures and crypt-like proliferative zones that expressed intestinal stem cell markers<sup>17</sup>. The epithelium contained functional enterocytes, as well as goblet, Paneth and enteroendocrine cells. Using this culture system as a model to study human intestinal development, we identified that the combined activity of WNT3A and FGF4 is required for hindgut specification whereas FGF4 alone is sufficient to promote hindgut morphogenesis. Our data indicate that human intestinal stem cells form *de novo* during development. We also determined that NEUROG3, a pro-endocrine transcription factor that is mutated in enteric anendocrinosis<sup>18</sup>, is both necessary and sufficient for human enteroendocrine cell development *in vitro*. PSC-derived human intestinal tissue should allow for unprecedented studies of human intestinal development and disease.

The epithelium of the intestine is derived from a simple sheet of cells called the definitive endoderm<sup>17</sup>. As a first step to generating intestinal tissue from PSCs (summarized in Supplementary Fig. 1), we used activin A, a nodal-related TGF- $\beta$  molecule, to promote differentiation into definitive endoderm as previously described<sup>11</sup>, resulting in up to 90% of the cells co-expressing the definitive endoderm markers SOX17 and FOXA2 and fewer than 2% expressing the mesoderm marker brachyury (Supplementary Fig. 2a). Using microarray analysis we observed a robust activation of definitive endoderm markers, many of which were expressed in mouse definitive endoderm from embryonic day (e)7.5 embryos (Supplementary Fig. 3 and Supplementary Table 1a, b). We investigated the intrinsic ability of definitive endoderm to form foregut and hindgut lineages by culturing for 7 days under permissive conditions and observed that cultures treated with activin A for only 3 days were competent to develop into both foregut (albumin (ALB)<sup>+</sup> and PDX1<sup>+</sup>) and hindgut (CDX2) lineages (Fig. 1b, control). In contrast, treatment with activin A for 4–5 days resulted in

definitive endoderm cultures that were intrinsically anterior in character and less competent in forming posterior lineages (Supplementary Fig. 2b).



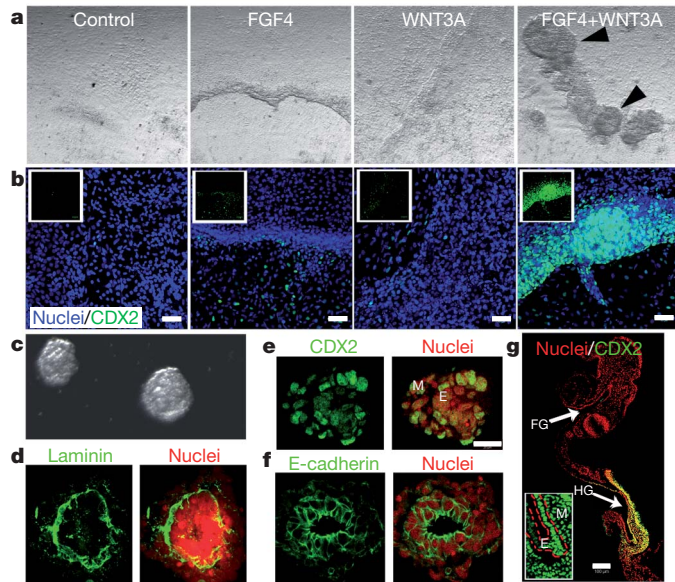
**Figure 1 | FGF4 and WNT3A act synergistically in a temporal and dose-dependent manner to specify stable posterior endoderm fate.** **a–d**, Activin A (100 ng ml<sup>-1</sup>) was used to differentiate H9 human ES cells into definitive endoderm. Definitive endoderm was treated with the posteriorizing factors FGF4 (50 or 500 ng), WNT3A (50 or 500 ng), or both for 6, 48 or 96 h. Cells were placed in permissive media for 7 days and expression of foregut markers (ALB, PDX1) and the hindgut marker (CDX2) were analysed by RT-qPCR (**a**) and immunofluorescence (**b–d**). The definitive endoderm of controls was grown for identical lengths of time in the absence of FGF4 or WNT3A. High levels of FGF4+WNT3A for 96 h resulted in stable CDX2 expression and lack of foregut marker expression. Scale bars, 50  $\mu$ m. Error bars are s.e.m. ( $n = 3$ ). \* $P < 0.05$ , \*\* $P < 0.001$ , \*\*\* $P < 0.0001$ .

<sup>1</sup>Division of Developmental Biology, Cincinnati Children's Hospital Medical Center, 3333 Burnet Avenue, Cincinnati, Ohio 45229-3039, USA. <sup>2</sup>Division of Gastroenterology, Hepatology and Nutrition, Cincinnati Children's Hospital Medical Center, 3333 Burnet Avenue, Cincinnati, Ohio 45229-3039, USA. <sup>3</sup>Division of Hematology and Oncology Cincinnati Children's Hospital Medical Center, 3333 Burnet Avenue, Cincinnati, Ohio 45229-3039, USA. <sup>4</sup>Division of Pulmonary Biology, Cincinnati Children's Hospital Medical Center, 3333 Burnet Avenue, Cincinnati, Ohio 45229-3039, USA.



Having identified the window of time when definitive endoderm fate was plastic (day 3 of activin A treatment), we used WNT3A and FGF4 to promote hindgut and intestinal specification. Studies in mouse, chick and frog embryos have demonstrated that Wnt and FGF signalling pathways are required for repressing anterior development and promoting posterior endoderm formation into the midgut and hindgut<sup>12–14</sup>. Consistent with this, conditioned media containing WNT3A was recently shown to promote *Cdx2* expression in mouse embryonic stem (ES)-cell-derived embryoid bodies<sup>19</sup>. In human definitive endoderm cultures, neither factor alone was sufficient to robustly promote a posterior fate (Supplementary Fig. 2c); but high concentrations of both FGF4 and WNT3A (FGF4+WNT3A) induced expression of the hindgut marker CDX2 in the definitive endoderm after 48 h (Supplementary Fig. 4). However, 48 h of FGF4+WNT3A treatment did not stably induce a CDX2<sup>+</sup> hindgut fate and expression of anterior markers PDX1 and albumin reappeared after cells were cultured in permissive media for 7 days (Fig. 1a, c). In contrast, 96 h of exposure to FGF4+WNT3A resulted in stable CDX2 expression and absence of anterior markers (Fig. 1a, d). These findings indicate a previously unidentified requirement for the synergistic activities of both the FGF and Wnt pathways in specifying the CDX2<sup>+</sup> mid/hindgut lineage.

Remarkably, FGF4+WNT3A-treated cultures underwent morphogenesis that was similar to embryonic hindgut formation. Between 2 and 5 days of FGF4+WNT3A treatment, flat cell sheets condensed into CDX2<sup>+</sup> epithelial tubes, many of which budded off to form floating hindgut spheroids (Fig. 2a–c, Supplementary Fig. 5a–f and Supplementary Table 2a). Spheroids were similar to e8.5 mouse hindgut and consisted of uniformly CDX2<sup>+</sup> polarized epithelium surrounded by CDX2<sup>+</sup> mesenchyme (Fig. 2d–g). Spheroids were completely devoid



**Figure 2 | Morphogenesis of posterior endoderm into three-dimensional, hindgut-like spheroids.** **a**, Bright-field images of definitive endoderm cultured for 96 h in media, FGF4, WNT3A or FGF4+WNT3A. FGF4+WNT3A cultures contained three-dimensional epithelial tubes and free-floating spheres (black arrows) **b**, CDX2 immunostaining (green) and nuclear stain (DRAQ5, blue) on cultures shown in **a**. Insets show CDX2 staining alone. **c**, Bright-field image of hindgut-like spheroids. **a–c**, Scale bars, 50  $\mu$ m. **d–f**, Analysis of CDX2, basal-lateral laminin and E-cadherin expression demonstrates an inner layer of polarized, cuboidal, CDX2<sup>+</sup> epithelium surrounded by non-polarized mesenchymal CDX2<sup>+</sup> cells. Scale bar in **e** is 20  $\mu$ m. **g**, CDX2 expression in an e8.5 mouse embryo (sagittal section). Inset is a magnified view showing that both hindgut endoderm (E, outlined with a red dashed line) and adjacent mesenchyme (M) are CDX2 positive (green). FG, foregut; HG, hindgut. Scale bar, 100  $\mu$ m.

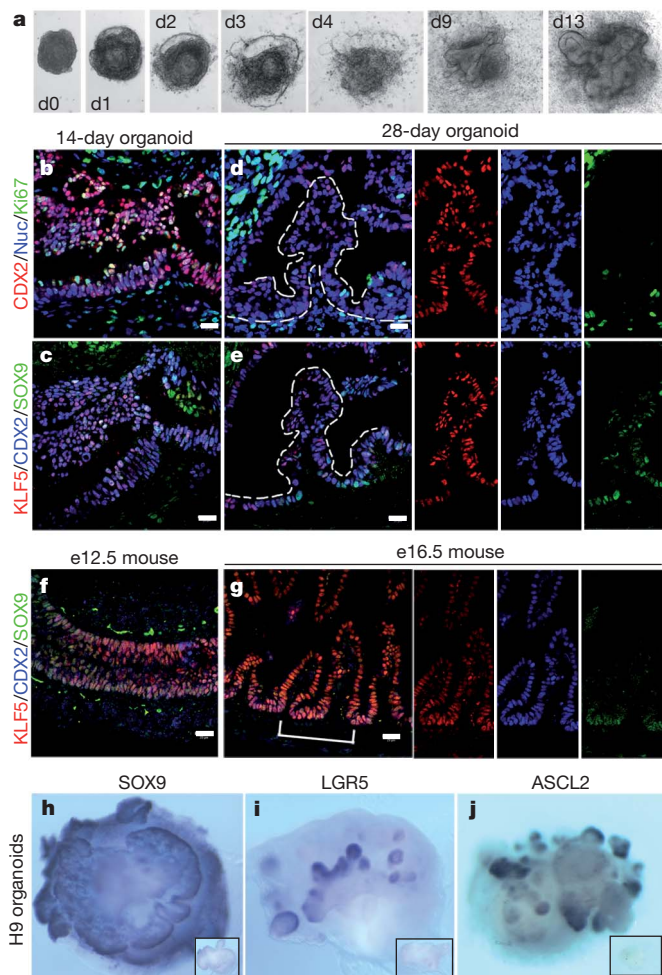
of albumin and PDX1-expressing foregut cells (Supplementary Fig. 5h, i). *In vitro* gut-tube morphogenesis was never observed in control or WNT3A-only treated cultures. FGF4-treated cultures had a twofold expansion of mesoderm and generated 4–10-fold fewer spheroids (Supplementary Fig. 2c and Supplementary Table 2a), which were weakly CDX2<sup>+</sup> and did not undergo further expansion (data not shown). Together our data support a mechanism for hindgut development where FGF4 promotes mesoderm expansion and morphogenesis, whereas FGF4 and WNT3A synergy is required for the specification of the hindgut lineage.

Importantly, this method for directed differentiation is broadly applicable to other PSC lines, as we were able to generate hindgut spheroids from both H1 and H9 human ES cell lines and from four induced PSC (iPSC) lines that we have generated and characterized (Supplementary Figs 3, 5 and 6). The kinetics of differentiation and the formation of spheroids were comparable between these lines (Supplementary Table 2). Two other iPSC lines tested were poor at hindgut spheroid formation and line iPSC3.6 also had a divergent transcriptional profile during definitive endoderm formation (Supplementary Fig. 3 and Supplementary Table 2c).

Whereas *in vivo* engraftment of PSC-derived cell types, such as pancreatic endocrine cells, has been used to promote maturation<sup>9</sup>, efficient development and maturation of organ tissues *in vitro* has proven more difficult. We investigated whether hindgut spheroids could develop and mature into intestinal tissue *in vitro* using recently described three-dimensional culture conditions that support growth and renewal of the adult intestinal epithelium<sup>15,16</sup>. When placed into this culture system, hindgut spheroids developed into intestinal organoids in a staged manner that was notably similar to fetal gut development (Fig. 3, Supplementary Fig. 5g and Supplementary Fig. 7). In the first 14 days the simple cuboidal epithelium of the spheroid expanded and formed a highly convoluted pseudostratified epithelium surrounded by mesenchymal cells (Fig. 3a–c), similar to an e12.5 fetal mouse gut (Fig. 3f). After 28 days, the epithelium matured into a columnar epithelium with villus-like involutions that protrude into the lumen of the organoid (Fig. 3d, e). Comparable transitions were observed during mouse fetal intestinal development (Fig. 3f, g and Supplementary Fig. 7). The spheroids expanded up to 40 fold in mass as they formed organoids (data not shown) and were split and passaged over 9 additional times and cultured for over 140 days with no signs of growth failure. The cellular gain during that time was up to 1,800 fold (data not shown), resulting in a total cellular expansion of 72,000 fold per hindgut spheroid. This directed differentiation was up to 50 fold more efficient than spontaneous embryoid body differentiation methods<sup>20</sup> (Supplementary Fig. 8) and resulted in organoids that were almost entirely intestinal (Supplementary Fig. 2e–g) as compared to embryoid bodies that contained a mix of neural, vascular and epidermal tissues (Supplementary Fig. 8).

Marker analysis showed that after 14 days in culture, virtually all of the epithelium expressed the intestinal transcription factors CDX2, KLF5 and SOX9 broadly and was highly proliferative (Fig. 3b, c). By 28 days, CDX2 and KLF5 remained broadly expressed in over 90% of the epithelium (Supplementary Fig. 2), whereas SOX9 became localized to pockets of proliferating cells at the base of the villus-like protrusions (Fig. 3d, e) similar to the intervillus epithelium of fetal mouse intestines at e16.5 (Fig. 3g and Supplementary Fig. 9). 5-bromodeoxyuridine (BrdU) pulse chase and analysis of organoids using a Z-stack series of confocal microscopic images showed that epithelial BrdU incorporation was largely restricted to SOX9-expressing cells in crypt-like structures that penetrated into the underlying mesenchyme (Supplementary Fig. 9). At 28 days, LGR5 is not expressed and ASCL2 (ref. 21) is broadly expressed and not restricted to the SOX9<sup>+</sup> proliferative zone. However, organoids cultured until 56 days expressed both ASCL2 and LGR5 in restricted epithelial domains that appear to overlap with the SOX9<sup>+</sup> zone (Fig. 3h–j and Supplementary Fig. 10). This domain is similar to developing intestinal progenitor domains *in vivo*,

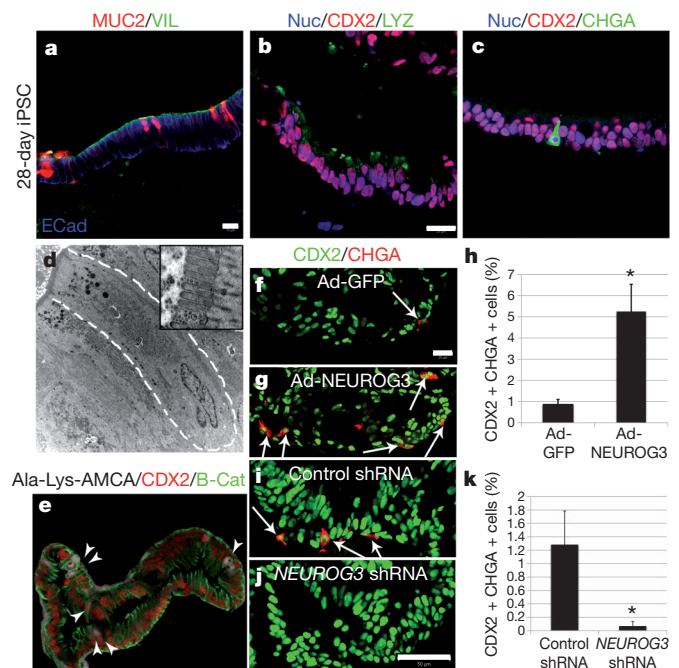




**Figure 3 | Human ES cells and iPSCs form three-dimensional intestine-like organoids.** **a**, A time course shows that intestinal organoids formed highly convoluted epithelial structures surrounded by mesenchyme after 13 days (**d**). **b–e**, Intestinal transcription factor expression (KLF5, CDX2, SOX9) and cell proliferation on serial sections of organoids after 14 and 28 days (serial sections are **b** and **c**, **d** and **e**). Ki67, nuclear proliferation antigen. Nuc, nuclei. **f, g**, Expression of KLF5, CDX2, and SOX9 in mouse fetal intestine at e14.5 (**f**) and e16.5 (**g**) is similar to developing intestinal organoids. The right panels show separate colour channels for **d**, **e** and **g** (bracket highlights the region shown in the panels on the right). **h–j**, Whole mount *in situ* hybridization of 56-day-old organoids showing epithelial expression of SOX9 (**h**) and restricted 'crypt-like' expression of the stem cell markers LGR5 (**i**) and ASCL2 (**j**). Insets show sense controls for each probe. Scale bars, 20  $\mu$ m.

which ultimately give rise to the stem cell niche in the crypt of Lieberkühn<sup>15</sup>. iPSCs were equally capable of forming intestinal progenitor domains (Supplementary Fig. 9e). Thus, PSC-derived intestinal epithelium continued to mature *in vitro* and develop proliferative domains with nascent intestinal stem cells.

Between 18 and 28 days in culture, we observed cytodifferentiation of the stratified epithelium into a columnar epithelium containing brush borders and all of the major cell lineages of the gut as determined by immunofluorescence and quantitative polymerase chain reaction with reverse transcription (RT-qPCR) (Fig. 4a–d and Supplementary Fig. 11). By 28 days of culture, villin (Fig. 4a) and DPPIV (not shown) were localized to the apical surface of the polarized columnar epithelium and transmission electron microscopy revealed a brush border of apical microvilli indistinguishable from those found in mature intestine (Fig. 4d and Supplementary Fig. 1). Enterocytes had a functional peptide transport system and were able to absorb a fluorescently



**Figure 4 | Formation and function of intestinal cell types and regulation of enteroendocrine differentiation by NEUROG3.** **a–c**, Twenty-eight-day iPSC-derived organoids were analysed for villin (VIL) and the goblet cell marker mucin (MUC2) (**a**), the Paneth cell marker lysozyme (LYZ) (**b**), or the endocrine cell marker chromogranin A (CHGA) (**c**). Nuc, nuclei. **d**, Electron micrograph showing an enterocyte cell with a characteristic brush border with microvilli (inset). **e**, Epithelial uptake of the fluorescently labelled dipeptide d-Ala-Lys-AMCA (arrowheads) indicating a functional peptide transport system. **f–h**, Adenoviral expression of NEUROG3 (Ad-NEUROG3) causes a fivefold increase in CHGA<sup>+</sup> cells compared to a GFP control (Ad-GFP). **n** = 4 biological samples; \**P* = 0.005. **i–k**, Organoids were generated from human ES cells that were stably transduced with shRNA-expressing lentiviral vectors. Compared to control shRNA organoids, NEUROG3 shRNA organoids had a 95% reduction in the number of CHGA<sup>+</sup> cells. **n** = 3 for shRNA controls and **n** = 5 for NEUROG3-shRNA; \**P* = 0.018. Scale bar in **a** is 10  $\mu$ m; all others are 20  $\mu$ m. Error bars are s.e.m.

labelled dipeptide (Fig. 4e)<sup>22</sup>. Cell counting revealed that the epithelium contained approximately 15% MUC2<sup>+</sup> goblet cells, which secrete mucins into the lumen of the organoid, 18% lysozyme-positive cells, which are indicative of Paneth cells, and ~1% chromogranin-A-expressing enteroendocrine cells (Fig. 4 and Supplementary Fig. 11g). MUC2 and lysozyme staining indicated that the goblet and Paneth cells in 28-day organoids are immature (Fig. 4a, b). However, in organoids that were passaged over 100 days, all cells had acquired a more mature phenotype and Paneth cells were often localized in crypt-like structures (Supplementary Fig. 12b, c). RT-qPCR confirmed the presence of additional markers of differentiated enterocytes (*FABP2*; also known as *IFABP*) and Paneth cells (*MMP7*) (Supplementary Fig. 11). Individual organoids seemed to be a mix of proximal intestine (GATA4<sup>+</sup>/GATA6<sup>+</sup>) and distal intestine (GATA4<sup>+</sup>/GATA6<sup>+</sup>; *HOXA13*-expressing) (Supplementary Figs 11 and 13)<sup>23</sup>. Thus, directed differentiation of PSCs into intestinal tissue *in vitro* is highly efficient in generating three-dimensional intestinal tissue containing crypt-like progenitor niches, villus-like domains and all of the differentiated cell types of the intestinal epithelium.

Intestinal organoids contained a mesenchymal layer that developed along with the epithelium in a staged manner similar to embryonic development<sup>10,24</sup> (Supplementary Fig. 14). Mesenchyme probably came from the 2% of mesoderm cells that were present after activin differentiation, which expanded up to 10% in FGF4-treated hindgut cultures (Supplementary Fig. 2). At 14 days, organoids broadly expressed mesenchymal markers including FOXF1 and vimentin

(Supplementary Fig. 14), similar to an e12.5 embryonic intestine (Supplementary Fig. 7). We also observed vimentin/smooth muscle actin (SMA; also known as ACTA2) double-positive cells indicative of intestinal subepithelial myofibroblasts<sup>25</sup>. By 28 days, we observed a layer of SMA<sup>+</sup>/desmin<sup>+</sup> double-positive cells, indicating smooth muscle, and desmin<sup>+</sup>/vimentin<sup>+</sup> fibroblasts<sup>26</sup>. The fact that intestinal mesenchyme differentiation coincided with differentiation of the overlying epithelium indicates that epithelial–mesenchymal crosstalk may be important in the development of PSC-derived intestinal organoids.

The molecular basis of congenital malformations in humans is often inferred from functional studies in model organisms. For example, neurogenin 3 (*NEUROG3*) was investigated as a candidate gene responsible for congenital loss of intestinal enteroendocrine cells in humans<sup>18</sup> because of its known role in enteroendocrine cell development in mouse<sup>27–30</sup>. However, it has been impossible to directly investigate the role of *NEUROG3* during human intestinal development. We therefore performed gain- and loss-of-function analyses to investigate the role of *NEUROG3* during human enteroendocrine cell development (Fig. 4 and Supplementary Fig. 15). *NEUROG3* was overexpressed in 28-day human organoids using adenoviral (Ad)-mediated transduction<sup>31</sup>. After 7 days, approximately 5% of cells were GFP<sup>+</sup> and Ad-*NEUROG3*–GFP-infected organoids contained fivefold more chromogranin A<sup>+</sup> endocrine cells than control organoids (Ad-enhanced GFP (eGFP)) (Fig. 4f–h and Supplementary Fig. 15), demonstrating that *NEUROG3* expression is sufficient to promote an enteroendocrine cell fate. To knockdown endogenous *NEUROG3*, we generated human ES cell lines by transducing cells with *NEUROG3* short hairpin (sh)RNA-expressing lentiviral vectors. *NEUROG3* mRNA levels were knocked down by 63% and this resulted in a 90% reduction in the number of enteroendocrine cells (Fig. 4i–k and Supplementary Fig. 15d–f), demonstrating that intestinal enteroendocrine cell development is highly dependent on *NEUROG3* expression. This indicates that partial loss-of-function mutations in human *NEUROG3* would be sufficient to cause a marked reduction in enteroendocrine cell numbers.

This is the first report, to our knowledge, demonstrating that human PSCs can be efficiently directed to differentiate *in vitro* into human tissue with a three-dimensional architecture and cellular composition remarkably similar to the fetal intestine. Moreover, PSC-derived human intestinal tissue undergoes maturation *in vitro*, developing intestinal stem cells and acquiring both absorptive and secretory functionality. This system allows for functional studies to investigate the molecular basis of human congenital gut defects *in vitro* and to generate intestinal tissue for eventual transplantation-based therapy for diseases such as necrotizing enterocolitis, inflammatory bowel diseases and short-gut syndromes. The ability to generate human intestinal tissues should also greatly facilitate future studies of intestinal stem cells and drug design to enhance absorption and bioavailability.

## METHODS SUMMARY

**Generation of human intestinal organoids.** Human ES cells and iPSCs were maintained on Matrigel (BD Biosciences) in mTesR1 medium without feeders. Differentiation into definitive endoderm was carried out as previously described<sup>11</sup>. Briefly, a 3-day activin A (R&D systems) differentiation protocol was used. Cells were treated with activin A (100 ng ml<sup>−1</sup>) for three consecutive days in RPMI 1640 medium (Invitrogen) with increasing concentrations of 0%, 0.2% and 2% HyClone defined fetal bovine serum (dFBS; Thermo Scientific). For hindgut differentiation, definitive endoderm cells were incubated in 2% dFBS–DMEM/F12 with 500 ng ml<sup>−1</sup> FGF4 and 500 ng ml<sup>−1</sup> WNT3A (R&D Systems) for up to 4 days. Between 2 and 4 days of treatment with growth factors, three-dimensional floating spheroids formed and were then transferred into three-dimensional cultures previously shown to promote intestinal growth and differentiation<sup>15,16</sup>. Briefly, spheroids were embedded in Matrigel (BD Bioscience) containing 500 ng ml<sup>−1</sup> R-Spondin1 (R&D Systems), 100 ng ml<sup>−1</sup> Noggin (R&D Systems) and 50 ng ml<sup>−1</sup> EGF (R&D Systems). After the Matrigel solidified, medium (advanced DMEM/F12; Invitrogen) supplemented with L-glutamine, 10 μM HEPES, N2 supplement (R&D Systems), B27 supplement (Invitrogen), and penicillin/streptomycin-containing growth factors was overlaid and replaced every 4 days.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 22 April; accepted 23 November 2010.**

**Published online 12 December 2010.**

- Mayhew, C. N. & Wells, J. M. Converting human pluripotent stem cells into  $\beta$ -cells: recent advances and future challenges. *Curr. Opin. Organ Transplant* **15**, 54–60 (2010).
- Spence, J. R. & Wells, J. M. Translational embryology: using embryonic principles to generate pancreatic endocrine cells from embryonic stem cells. *Dev. Dyn.* **236**, 3218–3227 (2007).
- Cai, J. *et al.* Directed differentiation of human embryonic stem cells into functional hepatic cells. *Hepatology* **45**, 1229–1239 (2007).
- D'Amour, K. A. *et al.* Production of pancreatic hormone-expressing endocrine cells from human embryonic stem cells. *Nature Biotechnol.* **24**, 1392–1401 (2006).
- Song, Z. *et al.* Efficient generation of hepatocyte-like cells from human induced pluripotent stem cells. *Cell Res.* **19**, 1233–1242 (2009).
- Zhang, D. *et al.* Highly efficient differentiation of human ES cells and iPSCs into mature pancreatic insulin-producing cells. *Cell Res.* **19**, 429–438 (2009).
- Basma, H. *et al.* Differentiation and transplantation of human embryonic stem cell-derived hepatocytes. *Gastroenterology* **136**, 990–999 (2008).
- Touboul, T. *et al.* Generation of functional hepatocytes from human embryonic stem cells under chemically defined conditions that recapitulate liver development. *Hepatology* **51**, 1754–1765 (2010).
- Kroon, E. *et al.* Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells *in vivo*. *Nature Biotechnol.* **26**, 443–452 (2008).
- Zorn, A. M. & Wells, J. M. Vertebrate endoderm development and organ formation. *Annu. Rev. Cell Dev. Biol.* **25**, 221–251 (2009).
- D'Amour, K. A. *et al.* Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nature Biotechnol.* **23**, 1534–1541 (2005).
- Dessimoz, J., Opoka, R., Kordich, J. J., Grapin-Botton, A. & Wells, J. M. FGF signaling is necessary for establishing gut tube domains along the anterior-posterior axis *in vivo*. *Mech. Dev.* **123**, 42–55 (2006).
- McLin, V. A., Rankin, S. A. & Zorn, A. M. Repression of Wnt/ $\beta$ -catenin signaling in the anterior endoderm is essential for liver and pancreas development. *Development* **134**, 2207–2217 (2007).
- Wells, J. M. & Melton, D. A. Early mouse endoderm is patterned by soluble factors from adjacent germ layers. *Development* **127**, 1563–1572 (2000).
- Gracz, A. D., Ramalingam, S. & Magness, S. T. Sox9 expression marks a subset of CD24-expressing small intestine epithelial stem cells that form organoids *in vitro*. *Am. J. Physiol. Gastrointest. Liver Physiol.* **298**, G590–G600 (2010).
- Sato, T. *et al.* Single Lgr5 stem cells build crypt-villus structures *in vitro* without a mesenchymal niche. *Nature* **459**, 262–265 (2009).
- de Santa Barbara, P., van den Brink, G. R. & Roberts, D. J. Development and differentiation of the intestinal epithelium. *Cell. Mol. Life Sci.* **60**, 1322–1332 (2003).
- Wang, J. *et al.* Mutant neurogenin-3 in congenital malabsorptive diarrhea. *N. Engl. J. Med.* **355**, 270–280 (2006).
- Cao, L. *et al.* Intestinal lineage commitment of embryonic stem cells. *Differentiation* doi:10.1016/j.diff.2010.09.182 (in the press).
- Torihashi, S. *et al.* Gut-like structures from mouse embryonic stem cells as an *in vitro* model for gut organogenesis preserving developmental potential after transplantation. *Stem Cells* **24**, 2618–2626 (2006).
- van der Flier, L. G. *et al.* Transcription factor achaete scute-like 2 controls intestinal stem cell fate. *Cell* **136**, 903–912 (2009).
- Groneberg, D. A., Doring, F., Eynott, P. R., Fischer, A. & Daniel, H. Intestinal peptide transport: ex vivo uptake studies and localization of peptide carrier PEPT1. *Am. J. Physiol. Gastrointest. Liver Physiol.* **281**, G697–G704 (2001).
- Haveri, H. *et al.* Transcription factors GATA-4 and GATA-6 in normal and neoplastic human gastrointestinal mucosa. *BMC Gastroenterol.* **8**, 9 (2008).
- McLin, V. A., Henning, S. J. & Jamrich, M. The role of the visceral mesoderm in the development of the gastrointestinal tract. *Gastroenterology* **136**, 2074–2091 (2009).
- Ormestad, M. *et al.* Foxf1 and Foxf2 control murine gut development by limiting mesenchymal Wnt signaling and promoting extracellular matrix production. *Development* **133**, 833–843 (2006).
- Kosinski, C. *et al.* Indian hedgehog regulates intestinal stem cell fate through epithelial–mesenchymal interactions during development. *Gastroenterology* **139**, 893–903 (2010).
- Jenny, M. *et al.* Neurogenin3 is differentially required for endocrine cell fate specification in the intestinal and gastric epithelium. *EMBO J.* **21**, 6338–6347 (2002).
- Lee, C. S., Perreault, N., Brestelli, J. E. & Kaestner, K. H. Neurogenin 3 is essential for the proper specification of gastric enteroendocrine cells and the maintenance of gastric epithelial cell identity. *Genes Dev.* **16**, 1488–1497 (2002).
- Lopez-Diaz, L. *et al.* Intestinal Neurogenin 3 directs differentiation of a bipotential secretory progenitor to endocrine cell rather than goblet cell fate. *Dev. Biol.* **309**, 298–305 (2007).
- Ootani, A. *et al.* Sustained *in vitro* intestinal epithelial culture within a Wnt-dependent stem cell niche. *Nature Med.* **15**, 701–706 (2009).
- Zhou, Q., Brown, J., Kanarek, A., Rajagopal, J. & Melton, D. A. *In vivo* reprogramming of adult pancreatic exocrine cells to  $\beta$ -cells. *Nature* **455**, 627–632 (2008).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank members of the laboratory, D. Wiginton and C. Wylie for input. We also thank M. Kofron, T. Stefader and R. Lang for assistance with imaging.

Vectors and antibodies were from D. Melton (Addgene no. 19410, 19413), S. Yamanaka (17217–17220), C. Baum (OCT4, KLF4, SOX4, MYC lenti), and I. Manabe (KLF5 antibody). This work was supported by the Juvenile Diabetes Research Foundation JDRF-2-2003-530 (J.M.W.) and NIH, R01GM072915 (J.M.W.); R01DK080823A1 and S1 (A.M.Z. and J.M.W.); R03 DK084167 and R01 CA142826 (N.F.S.), F32 DK83202-01 and T32 HD07463 (J.R.S.). We also acknowledge core support for viral vectors, microarrays (supported by P30 DK078392), karyotyping and the Pluripotent Stem Cell Facility (supported by U54 RR025216).

**Author Contributions** J.M.W. and J.R.S. conceived the study and experimental design, performed and analysed experiments and co-wrote the manuscript. S.A.R., M.F.K. and J.E.V.

performed experiments. C.N.M., M.F.K., K.T., V.V.K., J.E.V., E.E.H. and S.I.W. provided reagents, conceptual and/or technical support in generating and characterizing iPSC lines and intestinal organoids. N.F.S. and A.M.Z. provided additional conceptual and experimental support and co-funded the project. All authors read and approved the final manuscript.

**Author Information** Data have been deposited at NCBI under accession number GSE25557. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to J.M.W. ([james.wells@cchmc.org](mailto:james.wells@cchmc.org)).



## METHODS

**Maintenance of PSCs.** Human ES cells and induced pluripotent stem cells were maintained on Matrigel (BD Biosciences) in mTesR1 medium<sup>32,33</sup>. Cells were passaged approximately every 4 days, depending on colony density. To passage PSCs, they were washed with DMEM/F12 medium (no serum) (Invitrogen) and incubated in DMEM/F12 with 1 mg ml<sup>-1</sup> dispase (Invitrogen) until colony edges started to detach from the dish. The dish was then washed 3 times with DMEM/F12 medium. After the final wash, DMEM/F12 was replaced with mTesR1. Colonies were scraped off of the dish with a cell scraper and gently triturated into small clumps and passaged onto fresh Matrigel-coated plates.

**Differentiation of PSCs into definitive endoderm.** Differentiation into definitive endoderm was carried out as previously described<sup>11</sup>. Briefly, a 3-day activin A (R&D systems) differentiation protocol was used. Cells were treated with activin A (100 ng ml<sup>-1</sup>) for three consecutive days in RPMI 1640 media (Invitrogen) with increasing concentrations of 0%, 0.2% and 2% HyClone defined fetal bovine serum (dFBS; Thermo Scientific).

**Differentiation of definitive endoderm in permissive media.** After differentiation into definitive endoderm, cells were incubated in DMEM/F12 plus 2% dFBS with either 0, 50 or 500 ng ml<sup>-1</sup> FGF4 and/or 0, 50 or 500 ng ml<sup>-1</sup> WNT3A (R&D Systems) for 6, 48 or 96 h. Cultures were then grown in permissive medium consisting of DMEM plus 10% FBS for an additional 7 days.

**Directed differentiation into hindgut and intestinal organoids.** After differentiation into definitive endoderm, cells were incubated in 2% dFBS-DMEM/F12 with either 50 or 500 ng ml<sup>-1</sup> FGF4 and/or 50 or 500 ng ml<sup>-1</sup> WNT3A (R&D Systems) for 2–4 days. After 2 days with treatment of growth factors, three-dimensional floating spheroids were present in the culture. Three-dimensional spheroids were transferred into an *in vitro* system to support intestinal growth and differentiation previously described<sup>15,16</sup>. Briefly, spheroids were embedded in Matrigel (BD Bioscience; no. 356237) containing 500 ng ml<sup>-1</sup> R-Spondin1 (R&D Systems), 100 ng ml<sup>-1</sup> Noggin (R&D Systems) and 50 ng ml<sup>-1</sup> EGF (R&D Systems). After the Matrigel solidified, medium (advanced DMEM/F12; Invitrogen) supplemented with L-glutamine, 10 µM HEPES, N2 supplement (R&D Systems), B27 supplement (Invitrogen), and penicillin/streptomycin-containing growth factors was overlaid and replaced every 4 days.

**Generation and characterization of iPSC lines.** Normal human skin keratinocytes (HSKs) were obtained from donors with informed consent (Cincinnati Children's Hospital Medical Center (CCHMC) Institutional Review Board protocol CR1\_2008-0899). Normal HSKs were isolated from punch biopsies following trypsinization and subsequent culture on irradiated NIH3T3 feeder cells in F medium<sup>34</sup>. For iPSC generation, normal HSKs were transduced on two consecutive days with a 1:1:1:1 mix of recombinant RD114-pseudotyped retroviruses expressing OCT4, SOX2, KLF4 and MYC<sup>35,36</sup> in the presence of 8 µg ml<sup>-1</sup> polybrene. Twenty-four hours after the second transduction the virus mix was replaced with fresh F medium and cells were incubated for an additional three days. Cells were then trypsinized and seeded into 6-well dishes containing 1.875 × 10<sup>5</sup> irradiated mouse fibroblasts per well and Epilife medium. On the following day, medium was replaced with DMEM/F12 50:50 medium supplemented with 20% knockout serum replacement, 1 mM L-glutamine, 0.1 mM β-mercaptoethanol, 1 × non-essential amino acids, 4 ng ml<sup>-1</sup> basic fibroblast growth factor, and 0.5 mM valproic acid. Morphologically identifiable iPSC colonies arose after 2–3 weeks and were picked manually, expanded and analysed for expression of human PSC markers NANOG, DNMT3B, and using the antigen antibodies Tra1-60 and Tra1-81<sup>37,38</sup>. Early passage iPSC lines were adapted to feeder-free culture conditions consisting of maintenance in mTesR1 (Stem Cell Technologies) in culture dishes coated with Matrigel (BD Biosciences) and lines were karyotyped.

**Microarray analysis of human ES cells, iPSCs and definitive endoderm cultures.** For microarray analysis, RNA was isolated from undifferentiated and 3-day activin-treated human ES cell and iPSC cultures and used to create target DNA for hybridization to Affymetrix Human 1.0 Gene ST Arrays using standard procedures (Affymetrix). Independent biological triplicates were performed for each cell line and condition. Affymetrix microarray Cel files were subjected to RMA normalization in GeneSpring 10.1. Probe sets were first filtered for those that are overexpressed or underexpressed and then subjected to statistical analysis for differential expression by 2 fold or more between undifferentiated and differentiated cultures with *P* < 0.05 using the Students *t*-test. Log2 gene expression ratios

were then subjected to hierarchical clustering using the standard correlation distance metric as implemented in GeneSpring.

**Adenoviral-mediated expression of NEUROG3.** Adenoviral plasmids were obtained from Addgene and particles were generated as previously described<sup>31</sup>. Transduction was done on 28-day organoids that were removed from Matrigel, manually bisected then incubated in Ad-GFP or Ad-NEUROG3 viral supernatant and medium at a 1:1 ratio for approximately 4 h. Organoids were then re-embedded in Matrigel and incubated overnight with viral supernatant and medium at a 1:1 ratio. The next day, fresh organoid medium was placed on the cultures and was changed as described until the end of the experiment.

**shRNA knockdown human ES cell lines.** GpZ shRNA lentiviral vectors were obtained from Open Biosystems (GpZ-NEUROG3 Open Biosystems clone no. v2lhs\_309089; v2lhs\_309091; v2lhs\_309093; v2lhs\_309092 and GpZ-Control; Openbiosystems clone no. RHS4346). The CCHMC Viral Vector Core produced high-titre lentiviral particles for each plasmid. Low-passage H9 human ES cells were dissociated into a single-cell suspension using Accutase, were spun down and resuspended in mTesR1 containing 10 µM Y-27632. Cells were plated at low density and incubated with lentivirus for 24 h. For the NEUROG3 shRNA knockdown line, particles from all four vectors were used. mTesR1 was replaced daily, and after 72 h selection for puromycin- (2–4 µg ml<sup>-1</sup>) resistant human ES cells was carried out. Puromycin-resistant colonies were routinely maintained and passaged in mTesR1 + puromycin (4 µg ml<sup>-1</sup>).

**β-Ala-Lys-AMCA uptake.** β-Ala-Lys-AMCA was purchased from BioTrend Chemicals and was resuspended in water. Intestinal organoids were cut in half using a scalpel and were incubated for four hours in advanced DMEM/F12 plus 24 µM β-Ala-Lys-AMCA. Following incubation, organoids were washed several times in PBS, embedded in OCT freezing medium and were frozen at –70 °C. Ten-micrometre cryosections were cut and processed for standard immunohistochemistry.

**Tissue processing, immunohistochemistry and microscopy.** Tissues were fixed for 1 h to overnight in 4% paraformaldehyde or 3% glutaraldehyde for transmission electron microscopy (TEM). Cultured PSCs and definitive endoderm cells were stained directly. Hindgut and intestinal organoids were embedded in paraffin, epoxy resin LX-112 (Ladd Research), or frozen in OCT. Sections were cut at 6–10 µm for standard microscopy and 0.1 µm for TEM. TEM sections were stained with uranyl acetate. Paraffin sections were deparaffinized, subjected to antigen retrieval, blocked in the appropriate serum (5% serum in 1 × PBS plus 0.5% Triton-X) for 30 min, and incubated with primary antibody overnight at 4 °C. Slides were washed and incubated in secondary antibody in blocking buffer for 2 h at room temperature (23 °C). For a list of antibodies used and dilutions, see Supplementary Table 3. Slides were washed and mounted using Fluoromount-G. Confocal images were captured on a Zeiss LSM510 and Z-stacks were analysed and assembled using AxioVision software. An Hitachi H7600 transmission electron microscope was used to capture images.

**RNA isolation, RT-qPCR.** RNA was isolated using the Nucleospin II RNA isolation kit (Clontech). Reverse transcription was carried out using the SuperScriptIII Superscript (Invitrogen) according to manufacturer's protocol. Finally, qPCR was carried out using Quantitect SybrGreen MasterMix (Qiagen) on a Chromo4 Real-Time PCR (BioRad). PCR primers sequences were typically obtained from qPrimerDepot (<http://primerdepot.nci.nih.gov/>). Primer sequences are available upon request.

32. Ludwig, T. E. *et al.* Feeder-independent culture of human embryonic stem cells. *Nature Methods* **3**, 637–646 (2006).
33. Ludwig, T. E. *et al.* Derivation of human embryonic stem cells in defined conditions. *Nature Biotechnol.* **24**, 185–187 (2006).
34. Lambert, P. F. *et al.* Using an immortalized cell line to study the HPV life cycle in organotypic "raft" cultures. *Methods Mol. Med.* **119**, 141–155 (2005).
35. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).
36. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
37. Richards, M., Tan, S. P., Tan, J. H., Chan, W. K. & Bongso, A. The transcriptome profile of human embryonic stem cells as defined by SAGE. *Stem Cells* **22**, 51–64 (2004).
38. Thomson, J. A. *et al.* Embryonic stem cell lines derived from human blastocysts. *Science* **282**, 1145–1147 (1998).

# A unique chromatin signature uncovers early developmental enhancers in humans

Alvaro Rada-Iglesias<sup>1</sup>, Ruchi Bajpai<sup>1</sup>, Tomek Swigut<sup>1</sup>, Samantha A. Brugmann<sup>1</sup>, Ryan A. Flynn<sup>1</sup> & Joanna Wysocka<sup>1,2</sup>

Cell-fate transitions involve the integration of genomic information encoded by regulatory elements, such as enhancers, with the cellular environment<sup>1,2</sup>. However, identification of genomic sequences that control human embryonic development represents a formidable challenge<sup>3</sup>. Here we show that in human embryonic stem cells (hESCs), unique chromatin signatures identify two distinct classes of genomic elements, both of which are marked by the presence of chromatin regulators p300 and BRG1, monomethylation of histone H3 at lysine 4 (H3K4me1), and low nucleosomal density. In addition, elements of the first class are distinguished by the acetylation of histone H3 at lysine 27 (H3K27ac), overlap with previously characterized hESC enhancers, and are located proximally to genes expressed in hESCs and the epiblast. In contrast, elements of the second class, which we term 'poised enhancers', are distinguished by the absence of H3K27ac, enrichment of histone H3 lysine 27 trimethylation (H3K27me3), and are linked to genes inactive in hESCs and instead are involved in orchestrating early steps in embryogenesis, such as gastrulation, mesoderm formation and neurulation. Consistent with the poised identity, during differentiation of hESCs to neuroepithelium, a neuroectoderm-specific subset of poised enhancers acquires a chromatin signature associated with active enhancers. When assayed in zebrafish embryos, poised enhancers are able to direct cell-type and stage-specific expression characteristic of their proximal developmental gene, even in the absence of sequence conservation in the fish genome. Our data demonstrate that early developmental enhancers are epigenetically pre-marked in hESCs and indicate an unappreciated role of H3K27me3 at distal regulatory elements. Moreover, the wealth of new regulatory sequences identified here provides an invaluable resource for studies and isolation of transient, rare cell populations representing early stages of human embryogenesis.

Recent reports demonstrated that active enhancers can be identified by epigenomic profiling of p300 (ref. 4), H3K4me1 and H3K27ac<sup>5,6</sup>. To characterize the enhancer repertoire of hESCs we performed chromatin immunoprecipitation coupled to massively parallel DNA sequencing (ChIP-seq) using antibodies recognizing chromatin regulators (that is, p300, BRG1) and histone modifications (that is, H3K4me1, H3K27ac, H3K4me3, H3K27me3) that distinguish distal elements from proximal promoters<sup>5,6</sup> (Supplementary Fig. 1). As expected, previously characterized hESC enhancers (for example, *NANOG* (ref. 7) and *OCT4* (also called *POU5F1*)<sup>8</sup>) were bound by p300 and flanked by H3K4me1 and H3K27ac marked chromatin, but were not enriched for H3K27me3 or H3K4me3 (Fig. 1a and Supplementary Fig. 2a). Genome-wide analysis defined 5,118 genomic regions (hereafter referred to as class I elements) marked by a similar chromatin signature (that is, high p300, H3K4me1 and H3K27ac, low, if any, H3K4me3, and absence of H3K27me3), representing putative active hESC enhancers (Fig. 1b and Supplementary Data 1).

Interestingly, in the vicinity of many early developmental genes we noted promoter-distal p300-bound regions that were marked by

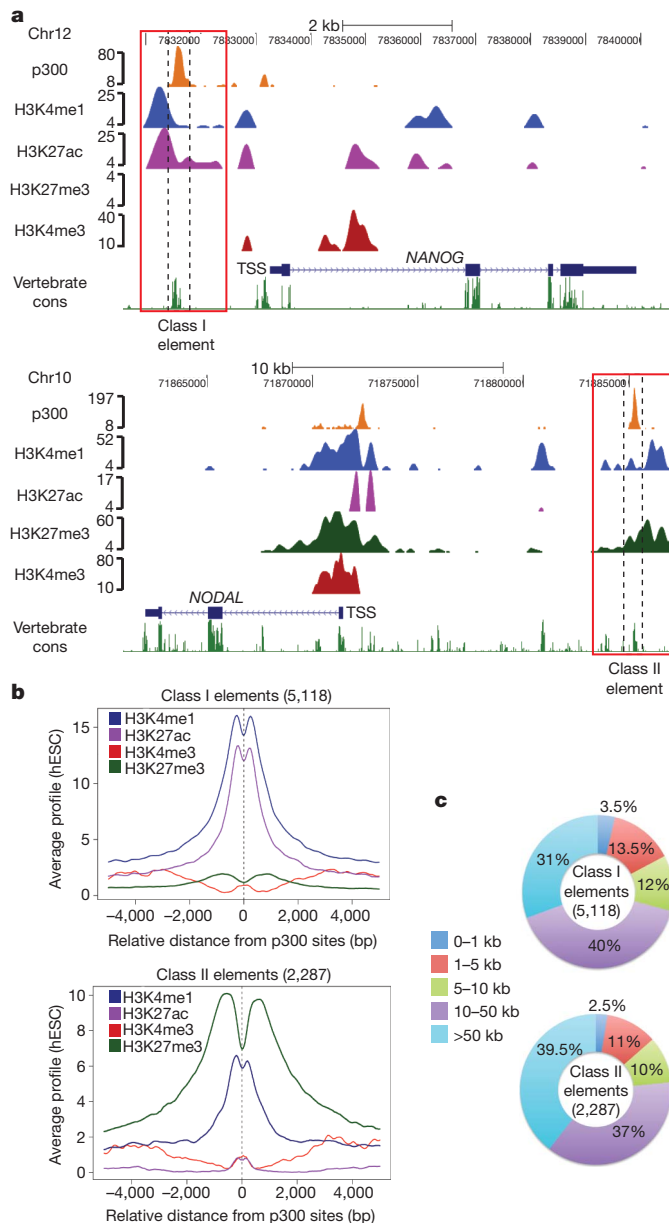
H3K4me1 but, in contrast to the active hESC enhancers, lacked H3K27ac and were instead enriched for H3K27me3, a modification associated with polycomb silencing<sup>9</sup> (Fig. 1a). Overall, we identified 2,287 p300-bound regions devoid of H3K27ac and marked by H3K27me3, which we will hereafter refer to as class II elements (Fig. 1b and Supplementary Data 1). In general, class II elements showed enrichment of both H3K27me3 and H3K4me1 flanking p300 peaks (Fig. 1b). In contrast, analysis of previously described adult tissue-specific enhancers<sup>10–13</sup> revealed no enrichment for any of the interrogated modifications (Supplementary Fig. 2b–e).

p300 enrichment levels were comparable at class I and II elements (Supplementary Fig. 3a), both classes were bound by BRG1 (Supplementary Fig. 3b), and showed similar genomic distribution relative to annotated transcription start sites (TSS), with over 95% of regions located away from promoters (Fig. 1c). Moreover, only 1.7% and 3.9% of class I and class II elements, respectively, overlapped with CpG islands, in sharp contrast to the 50% overlap observed for promoters. Another property of enhancers is their relative nucleosomal depletion compared to the flanking regions<sup>14,15</sup>. Using FAIRE-seq (formaldehyde-assisted isolation of regulatory elements<sup>16</sup> coupled to sequencing) we showed that class I and II elements were comparably nucleosome-depleted (Supplementary Fig. 3c). Furthermore, examination of a reported DNA-methylation-sensitive restriction enzyme data set from hESCs<sup>17</sup> revealed similar levels of DNA hypomethylation at class I and class II elements (Supplementary Fig. 3d).

ChIP-seq results were validated by ChIP-qPCR at a representative subset of class I and class II elements (labelled with the name of their closest gene) (Supplementary Figs 4a–d and 5). Further examination of the H3K27ac and H3K27me3 enrichments showed a mutually exclusive marking pattern at class I and class II elements (Supplementary Fig. 6). Sequential ChIP-qPCR demonstrated a simultaneous presence of H3K4me1/K27ac at class I regions, and H3K4me1/K27me3 at class II regions, indicating that the concurrent enrichments of H3K4me1 and H3K27me3 were not due to cell population heterogeneity (Fig. 2a, b). Moreover, consistent with H3K27me3, we observed enrichment of the PRC2 component, SUZ12, at class II elements (Supplementary Fig. 4e). We also detected preferential association of RNA POL2 with class I elements, as compared to class II elements, including its unphosphorylated, Ser5 phosphorylated and Ser 2 phosphorylated forms (Supplementary Fig. 7a–c).

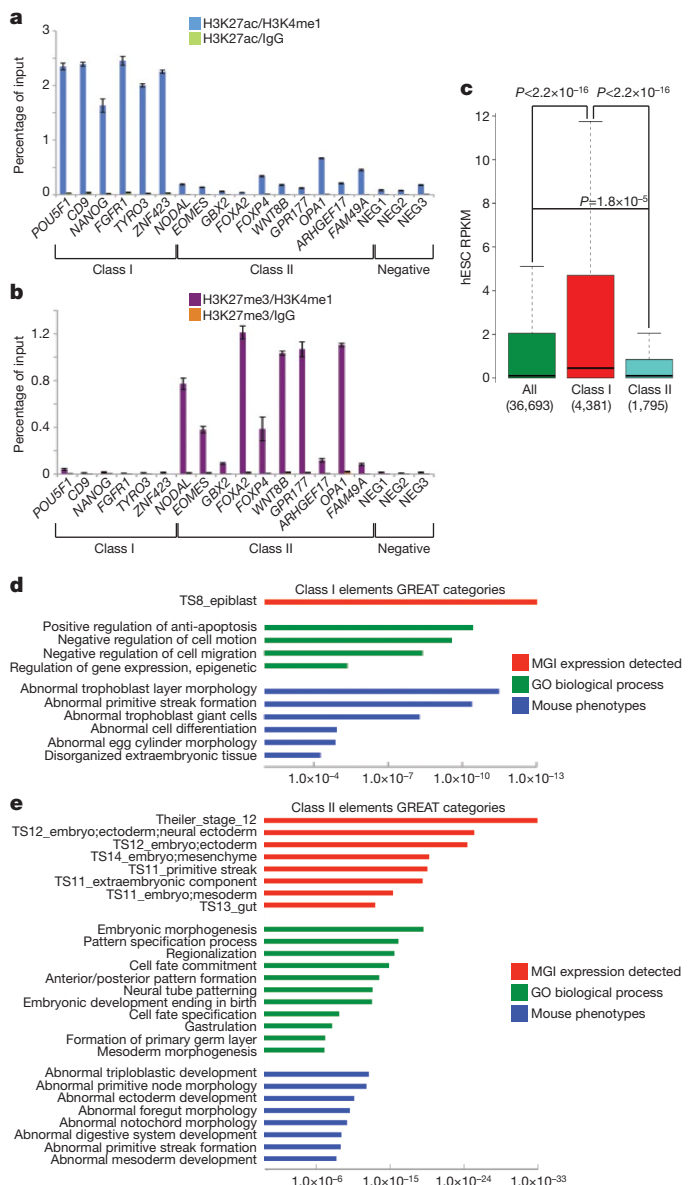
Next we asked whether transcriptional status of nearby genes differs between the two classes. To this end, we analysed hESC transcriptome by RNA-seq and examined transcripts originating from TSS closest to the elements of each class. Class-I-associated gene expression was significantly higher than expression of all genes, or of class-II-associated genes, which were poorly expressed (Fig. 2c). In agreement, class-II-associated TSS were enriched for both H3K27me3 and H3K4me3, whereas class-I-associated TSS were marked by high H3K4me3 levels (Supplementary Fig. 8a, c). Thus, the two classes defined by unique chromatin signatures are also distinguished by the transcriptional status of associated genes.

<sup>1</sup>Department of Chemical and Systems Biology, Stanford University School of Medicine, Stanford, California 94305, USA. <sup>2</sup>Department of Developmental Biology, Stanford University School of Medicine, Stanford, California 94305, USA.



**Figure 1 | Unique chromatin signatures distinguish two classes of regulatory elements in hESCs.** **a**, Genome browser representations of p300, H3K4me1, H3K27ac, H3K27me3 and H3K4me3 enrichment profiles in hESCs are shown for a representative class I (for example, *NANOG*, top) and class II (for example, *NODAL*, bottom) element and its flanking regions. The peak height corresponds to normalized fold enrichments as calculated by QuEST. **b**, Average hESC ChIP-seq signal profiles were generated for the indicated histone modifications around the central position of p300-bound regions, over class I (top) and class II (bottom) elements, respectively. **c**, Class I and II elements were mapped to their closest Ensembl gene TSS and the distribution of distances between elements and TSS is shown.

To investigate whether the two classes are linked to genes of distinct functional annotations, we performed ontology analysis using the Genomic Regions Enrichment of Annotations Tool (GREAT)<sup>18</sup> (Fig. 2d, e and Supplementary Data 2 and 3). Class I elements showed association with genes expressed in the epiblast, whose mouse homologues exhibit knockout phenotypes with defects in pre- and peri-implantation development (Fig. 2d). In contrast, class II elements are linked to genes expressed at, and essential for, gastrulation, germ-layer formation, neurulation and early somitogenesis (including *NODAL*, *EOMES*, *LEFTY2*, *EN1*, as well as *FOX*, *SOX* and *WNT* family members) (Fig. 2e). Notably, we did not observe enrichment of adult-tissue



**Figure 2 | Functional and molecular characterization of class I and II elements.** **a**, **b**, Sequential ChIP experiments were performed from hESCs with the indicated pairs of histone modification antibodies. ChIP material was analysed by qPCR for select class I and class II elements, as well as negative control regions (NEG1–3). The y axis shows per cent input recovery; error bars represent standard deviation (s.d.) from three technical replicates. **c**, RNA-seq data set was obtained from hESC poly(A)-RNA and reads per kilobase per million mapped reads (RPKM) were calculated for all human Ensembl genes. RPKMs for all annotated genes (green) or for those closest to class I (red) or class II (blue) elements are represented as box plots. *P*-values were calculated using non-paired Wilcoxon tests. In the box plots, bottom and top of the boxes correspond to the 25th and 75th percentiles and the internal band is the 50th percentile (median). The plot whiskers extending outside the boxes correspond to the lowest and highest datum within 1.5 interquartile range of the lower and upper quartiles, respectively. **d**, **e**, Functional annotation of class I (**d**) and class II (**e**) elements was performed using GREAT. The top over-represented categories belonging to three different ontologies are shown: Mouse Genome Informatics (MGI) expression detected (red) contains information on tissue- and developmental-stage-specific expression in mouse; Gene Ontology (GO) biological process (green) describes the biological processes associated with gene function; mouse phenotypes (blue) ontology contains data about mouse genotype–phenotype associations. The x axes values (in logarithmic scale) correspond to the binomial raw (uncorrected) *P*-values.

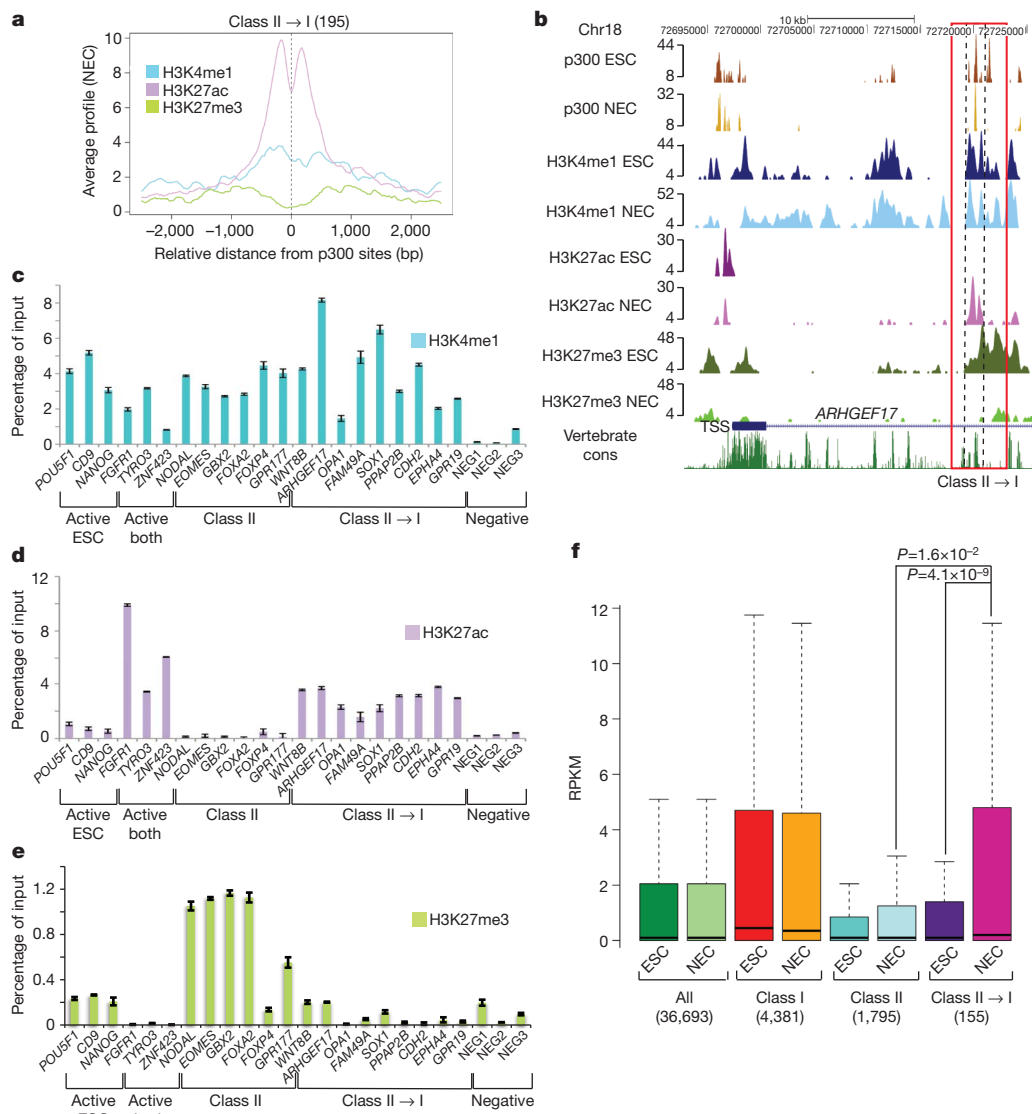


categories among class-II-linked genes, indicating no association with late enhancers.

Taken together, our results suggest that class II elements represent poised enhancers, which reveal their cell-type-dependent activity during development. One prediction from this hypothesis is that upon differentiation to a specific fate, a subset of poised enhancers linked to genes induced in this fate should acquire an active, class I signature. To test this prediction, we differentiated hESCs into neuroectodermal spheres (hNECs)<sup>19</sup>, generated p300, H3K4me1, H3K27ac and H3K27me3 profiles by ChIP-seq, and identified genomic elements that were marked by class II signature in hESCs, but acquired a strong enrichment of H3K27ac in hNECs (195 unique regions, Supplementary Data 1). Histone modification profiling over these regions showed concomitant decrease in H3K27me3 (Fig. 3a, b and Supplementary Fig. 9a) and we

refer to them hereafter as class II→I elements. Of note, a large number of the remaining class II regions (that is, those that did not acquire H3K27ac) retained H3K4me1 and H3K27me3 signature in hNECs, but showed diminished p300 occupancy (Supplementary Fig. 9b–d).

The aforementioned observations were validated by ChIP-qPCR for a representative subset of enhancers (Fig. 3c–e). We further showed that class II→I elements acquired RNA POL2 enrichment in hNECs, whereas hESC-specific active enhancers showed diminished RNA POL2 binding (Supplementary Fig. 10a). In agreement with a report documenting short bidirectional transcripts originating from enhancers<sup>20</sup>, we detected an increased level of bidirectional transcription from class II→I elements upon differentiation to hNECs, whereas transcripts originating from *NANOG* and *OCT4* enhancers were downregulated (Supplementary Fig. 10b, c).



**Figure 3 | A subset of class II elements acquires active enhancer chromatin signature upon neuroectodermal differentiation.** **a**, Average hNEC ChIP-seq signal profiles were generated for the indicated histone modifications around the central position of those p300-bound regions (as determined in hESC) that acquired H3K27ac enrichment in hNECs (that is, class II→I elements). **b**, Genome browser representation of p300, H3K4me1, H3K27ac and H3K27me3 (in hESCs and hNECs) binding profiles at a representative class II→I element. The peak height corresponds to normalized fold enrichments as calculated by QuEST. **c–e**, ChIP-qPCR analyses from hNECs with indicated histone modification antibodies at select elements including: class I elements that were only active in hESCs (active ESC), or in both hESCs and hNECs

(active both), or class II elements that did not acquire H3K27ac in hNEC (class II), or class II→I elements. The y axis shows per cent input recovery; error bars represent s.d. from three technical replicates. ChIPs used in these qPCRs represent biological replicates of those samples used in ChIP-seq. **f**, RNA-seq data sets from hESC and hNEC poly(A)-RNA were used to calculate the RPKM for all human Ensembl genes. RPKMs in both cell types are represented as box plots for all genes (All), genes linked to class I elements, genes linked to class II elements, and genes linked to class II→I elements. P-values were calculated using paired (NEC class II→I versus ESC class II→I) or non-paired (NEC class II→I versus NEC class II) Wilcoxon tests.

GREAT annotation of class II→I elements showed association with genes expressed in neuroectoderm and related to abnormalities in nervous system development (Supplementary Fig. 11 and Supplementary Data 4). In agreement, hNEC RNA-seq transcriptome analysis revealed significant upregulation of class-II→I-associated genes upon differentiation, whereas expression of the remaining class-II-associated genes was persistently low (Fig. 3f). Moreover, H3K27me3 levels at class-II→I-associated TSS were diminished and H3K4me3 levels induced in hNECs as compared to hESCs, whereas modification profiles over TSS associated with the remaining class II elements were relatively unchanged (Supplementary Fig. 8b, d).

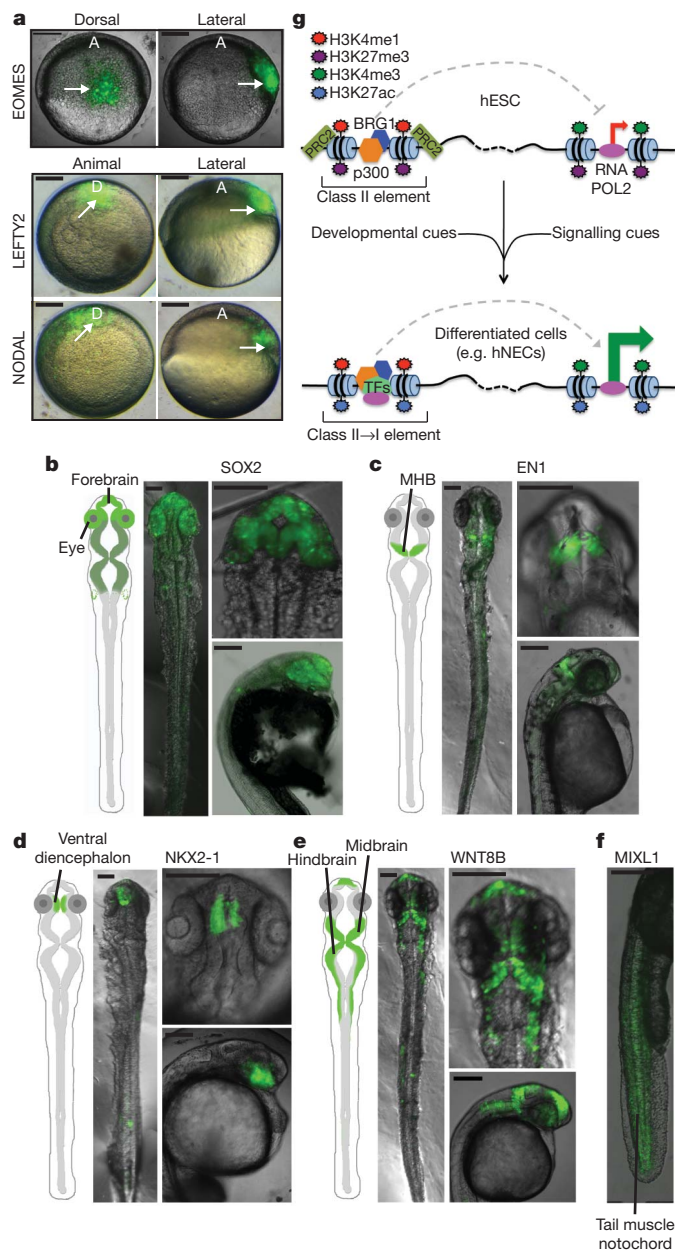
To examine if upon differentiation class II→I elements acquire the ability to drive gene expression, we infected hESCs with lentiviruses encoding a green fluorescent protein (GFP) reporter under the control of select class II→I (for example, *SOX2*, *HES1*), class I (for example, *CD9*, *JARID2*) and class II elements (for example, *EOMES*, *MYF5*) and monitored GFP fluorescence at day 1, 5 and 7 of differentiation to hNECs (Supplementary Table 1 and Supplementary Fig. 12). Class II→I reporters showed low, if any, fluorescence levels in hESCs, but were induced at day 5 of differentiation, whereas class I reporters

displayed a reverse pattern. Our results are consistent with class II elements representing poised developmental enhancers, which upon differentiation acquire, in a cell-type-dependent manner, the properties of active enhancers.

To test whether class II elements indeed function as developmental enhancers, we examined their activity during embryogenesis. Sequence conservation analysis revealed that class II elements are evolutionarily constrained and display a higher degree of conservation than class I elements (Supplementary Fig. 13a). VISTA enhancer browser search<sup>21</sup> identified fourteen class II elements for which enhancer activity was previously assayed at embryonic day 11.5 of mouse development. In nine cases, highly specific expression patterns were noted (Supplementary Table 2). Interestingly, two enhancers (*WNT8B*, *CDH2*) belong to the class II→I and, in agreement, drive gene expression specifically in neuroectoderm-derived structures in the mouse (Supplementary Table 2).

Next we screened enhancer activity of a select set of class II elements using zebrafish embryo transgenic reporter assay<sup>22,23</sup>. Selected elements correspond to previously uncharacterized human genomic sequences (except for *WNT8B*) that are located in proximity to genes whose zebrafish homologues have known expression patterns, although the elements themselves are generally not well conserved in the zebrafish genome (Supplementary Figs 13 and 14). GFP reporters were injected into one-cell-stage embryos and fluorescence was monitored throughout fish embryogenesis (Supplementary Fig. 15). For eight out of nine assayed class II reporters, specific and reproducible GFP patterns were observed at distinct developmental stages and anatomical locations (Fig. 4a–f, Supplementary Fig. 14 and Supplementary Table 3).

A first subgroup of assayed elements (for example, *NODAL*, *EOMES*, *LEFTY2*) drove gastrulation-specific expression at the shield, the fish equivalent of mouse primitive groove (Fig. 4a and Supplementary Fig. 14). Although none of the three tested sequences is well conserved in fish, proximal genes *NODAL*, *EOMES* and *LEFTY2* are conserved across vertebrates, with shield-specific expression pattern of zebrafish *NODAL* and *LEFTY2* homologues<sup>24</sup> (Supplementary Figs 14 and 16a). From mice to frogs, *EOMES* expression is initially restricted to the primitive groove and blastopore lip, respectively<sup>25,26</sup>, but the zebrafish *EOMES* homologue is only expressed at later stages (ZFIN database, identifier ZDB-PUB-051025-1). Remarkably, the element representing a putative *EOMES* enhancer drives shield-specific expression,



**Figure 4 | Class II elements have developmental enhancer activity *in vivo*.** **a**, Merged bright-field and GFP images are shown for representative shield stage zebrafish embryos injected with class II elements proximal to human *EOMES*, *LEFTY2* and *NODAL*. For the *EOMES* enhancer, dorsal (anterior to top) and lateral (shield to right) views are presented in the left and right panels, respectively. For *LEFTY2* and *NODAL*, animal pole (shield to top) and lateral (shield to right) views are presented in the left and right panels, respectively. White arrows indicate the location of the shield in each image. A, anterior; D, dorsal. Scale bar, 150  $\mu$ m. **b–f**, Merged bright-field and GFP images are shown for representative 24–28 h.p.f. zebrafish embryos injected with class II elements proximal to *SOX2* (**b**), *EN1* (**c**), *NKX2-1* (**d**), *WNT8B* (**e**) and *MIXL1* (**f**) genes. In **b–e**, schematics highlighting the relevant anatomical structures where GFP expression was reproducibly observed are shown on the left, and three images correspond, from left to right and top to bottom, to whole-embryo flattened dorsal views, dorsal anterior views and lateral anterior views, respectively. In **f**, a lateral posterior view is shown. In **b–f**, scale bar = 150  $\mu$ m. MHB, midbrain–hindbrain boundary. **g**, Proposed model for enhancer bookmarking during early embryonic development. Poised developmental enhancers (class II) are marked by a unique chromatin signature, involving occupancy of chromatin modifiers p300, BRG1 and PRC2 and nucleosomal regions marked by H3K4me1 and H3K27me3. During differentiation, appropriate developmental and signalling cues are able to rapidly transition these poised, pre-marked enhancers into an active state represented by the acquisition of H3K27ac, RNA POL2 binding, recruitment of tissue-specific transcription factors (TFs) and loss of H3K27me3, leading to the establishment of tissue-specific gene expression patterns.

indicating responsiveness of this human sequence to zebrafish gastrulation circuitry.

A second subgroup of class II reporters (for example, *SOX2*, *NKX2-1*, *EN1*, *WNT8B*, *MIXL1*) drove GFP expression at later developmental stages (24–28 h post fertilization (h.p.f.)) (Fig. 4b–f); this expression was restricted to specific anatomical structures such as the midbrain–hindbrain boundary (*EN1*)<sup>27</sup> or the ventral diencephalon/hypothalamus (*NKX2-1*)<sup>28</sup>. Again, despite the low degree of sequence conservation in fish (Supplementary Fig. 13), observed GFP patterns were generally consistent with the reported expression of the putative target gene homologues<sup>24,29</sup> (Supplementary Fig. 16b–d).

Importantly, specificity of our results was validated with an extensive set of control regions, including: (1) five class I elements; (2) four non-conserved genomic regions flanking select analysed class II elements; (3) four human adult tissue-specific enhancers; (4) three randomly selected intergenic non-conserved regions; (5) empty vector (Supplementary Table 4). All control regions showed only weak, diffused and nonspecific GFP patterns from 6 h.p.f. to 5 d.p.f. (Supplementary Figs 17–21). It is worth mentioning that based on our limited analysis, class I elements active in hESCs do not appear to drive pre-specification expression in zebrafish. Finally, to address whether expression patterns driven by class II elements are dynamic, we monitored several reporters (*LEFTY2*, *SOX2*, *EN1*, *NKX2-1*) throughout embryogenesis for up to 5 d.p.f. In all cases, GFP patterns were transient in nature, with fluorescence signals barely detectable after 3 d.p.f. (Supplementary Figs 17–21), further underscoring that class II regions represent dynamically regulated developmental enhancers.

We uncovered a unique chromatin signature that bookmarks early developmental enhancers in pluripotent cells, likely to prime them for a response to signalling and developmental cues (Fig. 4g). In addition to novel insights into gene regulation, our study identified a set of over 2,000 putative regulatory sequences, thereby creating an invaluable resource for lineage tracking and isolation of transient cell populations representing early steps of human development.

## METHODS SUMMARY

**ChIP-seq.** Approximately  $10^7$  hESCs or hNECs were used for each ChIP experiment. Cells were crosslinked with 1% formaldehyde for 10 min at 25 °C, chromatin was sonicated and immunoprecipitated with 3–5 µg of antibody. Sequencing libraries were prepared according to Illumina protocol from: hESC and hNEC p300 ChIP, hESC BRG1 ChIP, hESC FAIRE, hESC and hNEC H3K4me3 ChIP, hESC and hNEC H3K4me1 ChIPs, hESC and hNEC H3K27me3 ChIPs, hESC and hNEC H3K27ac ChIPs, hESC and hNEC input DNAs. Libraries were sequenced using Illumina Genome Analyser and resulting sequence reads mapped by ELAND (Illumina Inc.) and analysed by QuEST 2.4 (ref. 30).

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 4 August; accepted 25 November 2010.**

**Published online 15 December 2010.**

1. Bulger, M. & Groudine, M. Enhancers: the abundance and function of regulatory sequences beyond promoters. *Dev. Biol.* **339**, 250–257 (2010).
2. Hallikas, O. *et al.* Genome-wide prediction of mammalian enhancers based on analysis of transcription-factor binding affinity. *Cell* **124**, 47–59 (2006).
3. Visel, A., Rubin, E. M. & Pennacchio, L. A. Genomic views of distant-acting enhancers. *Nature* **461**, 199–205 (2009).
4. Visel, A. *et al.* ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* **457**, 854–858 (2009).
5. Heintzman, N. D. *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108–112 (2009).
6. Heintzman, N. D. *et al.* Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature Genet.* **39**, 311–318 (2007).
7. Chan, K. K. *et al.* KLF4 and PBX1 directly regulate NANOG expression in human embryonic stem cells. *Stem Cells* **27**, 2114–2125 (2009).
8. Yeom, Y. I. *et al.* Germ-line regulatory element of Oct-4 specific for the totipotent cycle of embryonic cells. *Development* **122**, 881–894 (1996).

9. Kerppola, T. K. Polycomb group complexes—many combinations, many functions. *Trends Cell Biol.* **19**, 692–704 (2009).
10. Cockerill, P. N. *et al.* Human granulocyte-macrophage colony-stimulating factor enhancer function is associated with cooperative interactions between AP-1 and NFATp/c. *Mol. Cell. Biol.* **15**, 2071–2079 (1995).
11. Nakabayashi, H. *et al.* Functional mapping of tissue-specific elements of the human  $\alpha$ -fetoprotein gene enhancer. *Biochem. Biophys. Res. Commun.* **318**, 773–785 (2004).
12. Itani, H. A., Liu, X., Pratt, J. H. & Sigmund, C. D. Functional characterization of polymorphisms in the kidney enhancer of the human renin gene. *Endocrinology* **148**, 1424–1430 (2007).
13. Segawa, K. *et al.* Identification of a novel distal enhancer in human adiponectin gene. *J. Endocrinol.* **200**, 107–116 (2009).
14. Mito, Y., Henikoff, J. G. & Henikoff, S. Histone replacement marks the boundaries of cis-regulatory domains. *Science* **315**, 1408–1411 (2007).
15. He, H. H. *et al.* Nucleosome dynamics define transcriptional enhancers. *Nature Genet.* **42**, 343–347 (2010).
16. Giresi, P. G. & Lieb, J. D. Isolation of active regulatory elements from eukaryotic chromatin using FAIRE (Formaldehyde Assisted Isolation of Regulatory Elements). *Methods* **48**, 233–239 (2009).
17. Harris, R. A. *et al.* Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nature Biotechnol.* **28**, 1097–1105 (2010).
18. McLean, C. Y. *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nature Biotechnol.* **28**, 495–501 (2010).
19. Bajpai, R. *et al.* Molecular stages of rapid and uniform neuralization of human embryonic stem cells. *Cell Death Differ.* **16**, 807–825 (2009).
20. Kim, T. K. *et al.* Widespread transcription at neuronal activity-regulated enhancers. *Nature* **465**, 182–187 (2010).
21. Visel, A., Minovitsky, S., Dubchak, I. & Pennacchio, L. A. VISTA Enhancer Browser—a database of tissue-specific human enhancers. *Nucleic Acids Res.* **35**, D88–D92 (2007).
22. Fisher, S. *et al.* Evaluating the biological relevance of putative enhancers using Tol2 transposon-mediated transgenesis in zebrafish. *Nature Protocols* **1**, 1297–1305 (2006).
23. Navratilova, P. *et al.* Systematic human/zebrafish comparative identification of cis-regulatory activity around vertebrate developmental transcription factor genes. *Dev. Biol.* **327**, 526–540 (2009).
24. Sprague, J. *et al.* The Zebrafish Information Network: the zebrafish model organism database. *Nucleic Acids Res.* **34**, D581–D585 (2006).
25. Hancock, S. N., Agulnik, S. I., Silver, L. M. & Papaioannou, V. E. Mapping and expression analysis of the mouse ortholog of *Xenopus Eomesodermin*. *Mech. Dev.* **81**, 205–208 (1999).
26. Ryan, K., Garrett, N., Mitchell, A. & Gurdon, J. B. Eomesodermin, a key early gene in *Xenopus* mesoderm differentiation. *Cell* **87**, 989–1000 (1996).
27. Danielian, P. S. & McMahon, A. P. Engrailed-1 as a target of the Wnt-1 signalling pathway in vertebrate midbrain development. *Nature* **383**, 332–334 (1996).
28. Marin, O., Baker, J., Puelles, L. & Rubenstein, J. L. Patterning of the basal telencephalon and hypothalamus is essential for guidance of cortical projections. *Development* **129**, 761–773 (2002).
29. Robb, L. *et al.* Cloning, expression analysis, and chromosomal localization of murine and human homologues of a *Xenopus* mix gene. *Dev. Dyn.* **219**, 497–504 (2000).
30. Valouev, A. *et al.* Genome-wide analysis of transcription factor binding sites based on ChIP-Seq data. *Nature Methods* **5**, 829–834 (2008).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank Wysocka laboratory members for ideas and manuscript comments; I. A. Shestopalov and J. K. Chen for sharing zebrafish resources, equipment and knowledge; T. Howes and D. M. Kingsley for the pT2HE vector; Z. Weng and A. Sidow for Illumina sequencing; and A. Valouev for discussion on ChIP-seq data analysis. This work was supported by WM Keck Foundation Distinguished Young Scholar in Biomedical Research Award and CIRRM RN1 00579-1 grant to J.W.A.R.-I. was supported by an EMBO long-term fellowship.

**Author Contributions** A.R.-I. conceived the project, performed and interpreted most experiments, including all genomic data analyses. R.B. established hESC culture and differentiation and performed most zebrafish imaging. T.S. generated enhancer reporter constructs, and together with S.A.B. and A.R.-I. participated in the *in vivo* enhancer screening. R.A.F. performed the RT-qPCR analysis of enhancer RNAs. J.W. contributed ideas and interpreted results. A.R.-I. and J.W. wrote the manuscript with input from all authors.

**Author Information** All sequencing data have been deposited in Gene Expression Omnibus (GEO) data repository under accession number GSE24447. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to J.W. ([wysocka@stanford.edu](mailto:wysocka@stanford.edu)).



## METHODS

**hESC culture.** hESCs (H9 line, Wi-Cell) were expanded in feeder-free, serum-free medium, mTESR-1 from StemCell technologies. Cells were passaged 1:7 every 5–6 days by incubation with accutase (Invitrogen) and resultant small cell clusters (50–200 cells) were subsequently re-plated on tissue culture dishes coated overnight with growth-factor-reduced matrigel (BD Biosciences). hESC quality was regularly tested by evaluating the expression of a panel of hESC markers (for example, alkaline phosphatase, OCT4) and the capacity to differentiate into cell types derived from the three germ layers.

**Neuroectoderm cell (NEC) differentiation.** hESCs were differentiated into hNECs using a previously described differentiation protocol<sup>21</sup>. Briefly, hESCs were incubated with 2 mg ml<sup>-1</sup> collagenase. Once detached, cells were plated in NEC differentiation media: 1:1 neurobasal medium/DMEM F-12 medium (Invitrogen), 0.5× B-27 supplement minus vitamin A (50× stock, Invitrogen), 0.5× N-2 supplement (100× stock, Invitrogen), 20 ng ml<sup>-1</sup> bFGF (Peprotech), 20 ng ml<sup>-1</sup> EGF (Sigma-Aldrich), 5 µg ml<sup>-1</sup> bovine insulin (Sigma-Aldrich), 0.1 µg ml<sup>-1</sup> recombinant human NOGGIN (Peprotech), 1× Glutamax-I supplement (100× stock, Invitrogen). Cells were differentiated for 7 days, changing media every other day.

**Chromatin immunoprecipitation (ChIP), sequential ChIP, FAIRE and antibodies.** ChIP assays were performed from approximately 10<sup>7</sup> hESCs or hNECs per experiment, according to previously described protocol with slight modifications<sup>31</sup>. Briefly, cells were crosslinked with 1% formaldehyde for 10 min at room temperature and formaldehyde was quenched by addition of glycine to a final concentration of 0.125 M. Chromatin was sonicated to an average size of 0.5–2 kb, using Bioruptor (Diagenode). A total of 3–5 µg of antibody was added to the sonicated chromatin and incubated overnight at 4 °C. 10% of chromatin used for each ChIP reaction was kept as input DNA. Subsequently, 75 µl of protein A or protein G Dynal magnetic beads (depending of antibody species and Ig isotype) were added to the ChIP reactions and incubated for four additional hours at 4 °C. Magnetic beads were washed and chromatin eluted, followed by reversal of the crosslinkings and DNA purification. Resultant ChIP DNA was dissolved in water.

Sequential ChIPs were performed as previously described with slight modifications<sup>32</sup>. Chromatin was prepared as described above for ChIP and after addition of the first antibody (3–5 µg) and corresponding washes, magnetic beads were resuspended in 75 µl TE/10 mM DTT. Samples were diluted 20 times with dilution buffer (1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8, 150 mM NaCl) and second antibody (3–5 µg) was added to each reaction. Beads were then washed, crosslinking reversed and DNA purified and dissolved in water.

For FAIRE, sonicated chromatin was prepared as for ChIP and DNA was extracted as previously described<sup>16</sup>.

All antibodies used in this study have been previously reported as suitable for ChIP: p300 (sc-585, Santa Cruz Biotechnology)<sup>5</sup>, BRG1 (clone JA1, a gift from G. Crabtree)<sup>33</sup>, H3K4me1 (ab8895, Abcam)<sup>5</sup>, H3K27ac (ab4729, Abcam)<sup>5</sup>, H3K4me3 (39159, Active Motif)<sup>34</sup>, H3K27me3 (39536, Active Motif)<sup>35</sup>, RNA POL2 unphosphorylated (8WG16 clone, MMS-126R, Covance)<sup>36</sup>, RNA POL2 ser5P (ab5131, Abcam)<sup>37</sup>, RNA POL2 ser2P (ab5095, Abcam)<sup>38</sup>, normal rabbit IgG (12-370, Millipore).

**ChIP-qPCR.** All primers used in qPCR analysis are shown in Supplementary Data 5. Primers are named after proximal putative target genes of the investigated enhancers. For each tested genomic element, two sets of primers were used, one set overlapping the peak of maximal p300 enrichment (central primers) and another set overlapping flanking regions with histone modification enrichments (flanking primers). This strategy was used because p300 peaks typically occurred within nucleosome-poor regions. qPCR analysis was performed in a Light Cycler 480II machine (Roche), using technical triplicates and ChIP-qPCR signals were calculated as percentage of input. Standard deviations were measured from the technical triplicate reactions and represented as error bars.

**RT-qPCR of enhancer RNAs.** To assess levels of enhancer-associated transcription, total RNA from hESCs and hNECs differentiated for 7 days was isolated using Trizol reagent followed by ethanol precipitation according to the manufacturer's protocol (Invitrogen). To remove genomic DNA contaminants, the Turbo DNA-Free kit was used following rigorous DNase treatment (two times, 30 min incubations at 37 °C). cDNA was generated from 100 ng of DNA-free RNA using the QuantiTect Reverse Transcription Kit (Qiagen) with two modifications: (1) The gDNA elimination reaction was extended for 5 min and (2) the reverse transcription elongation time was 30 min. Quantitative PCR (qPCR) primers were designed (Supplementary Data 5) to target regions surrounding the p300 peaks that defined each tested enhancer. qPCR runs and analysis were performed on the Light Cycler 480II machine (Roche). To calculate fold change between the hESCs and hNECs, the  $\Delta\Delta C_t$  method was used and the 18S rRNA transcripts were used as a loading control. Standard deviations were measured from technical triplicate reactions and were represented as error bars. Biological replicate experiments for hNECs were performed and very similar results were obtained (data not shown).

**ChIP-seq.** Libraries were prepared from: hESC and hNEC p300 ChIP, hESC BRG1 ChIP, hESC FAIRE, hESC and hNEC H3K4me3 ChIP, hESC and hNEC H3K4me1 ChIPs, hESC and hNEC H3K27me3 ChIPs, hESC and hNEC H3K27ac ChIPs, hESC and hNEC input DNAs. ChIP-seq, FAIRE-seq and input libraries were prepared according to Illumina protocol and sequenced using Illumina Genome Analyser. All sequences were mapped by ELAND software (Illumina Inc.) and analysed by QuEST 2.4 software<sup>30,35</sup>. ChIP-seq enrichment regions for the following profiled proteins were determined using the indicated settings, according to QuEST recommendations: hESC p300: KDE (kernel density estimation) bandwidth = 30, ChIP seeding fold enrichment = 30, ChIP extension fold enrichment = 3, ChIP-to-background fold enrichment = 3; hESC H3K4me3: KDE bandwidth = 60, ChIP seeding fold enrichment = 30, ChIP extension fold enrichment = 3, ChIP-to-background fold enrichment = 3; hESC H3K4me1: KDE bandwidth = 100, ChIP seeding fold enrichment = 10, ChIP extension fold enrichment = 3, ChIP-to-background fold enrichment = 2.5; hESC H3K27me3: KDE bandwidth = 100, ChIP seeding fold enrichment = 10, ChIP extension fold enrichment = 8, ChIP-to-background fold enrichment = 2.5; hESC and hNEC H3K27ac: KDE bandwidth = 100, ChIP seeding fold enrichment = 10, ChIP extension fold enrichment = 3, ChIP-to-background fold enrichment = 2.5.

For all ChIP-seq data sets, WIG files were generated with QuEST, which were subsequently used for visualization purposes and for obtaining average signal profiles.

**RNA-seq.** RNAs from hESCs and NECs were extracted with Trizol (Invitrogen), following the manufacturer's recommendations. 10 µg of total RNA were subjected to two rounds of oligo-dT purification using Dynal oligo-dT beads (Invitrogen). 100 ng of the purified RNA were fragmented with 10× fragmentation buffer (Ambion). Fragmented RNA was used for first-strand cDNA synthesis, using random hexamer primers (Invitrogen) and SuperScript II enzyme (Invitrogen). Second strand cDNA was obtained by adding RNaseH (Invitrogen) and DNA Pol I (New England Biolabs) to the first strand cDNA mix. The resulting double-stranded cDNA was used for Illumina library preparation as described for ChIP-seq experiments.

RNA-seq libraries were sequenced with Illumina Genome Analyser and both mapping and analysis of resulting reads were performed with DNAnexus software tools (<https://dnanexus.com>). Reads per kilobase per million mapped reads (RPKM) were calculated for all human Ensembl genes. The specificity and quality of our RNA-seq data can be visualized at several hESC- or hNEC-specific genes (Supplementary Fig. 22).

**Class I and class II element selection criteria.** ChIP-seq enrichment regions as determined by QuEST were used to define class I and class II elements (Supplementary Data 1). To this end, operations (intersection, subtraction, and so on) between genomic data sets were performed with GALAXY (<http://main.g2.bx.psu.edu/>) and the following selection criteria were used: class I elements (5,518 regions): genomic regions with hESC p300 enrichment (ChIP seeding fold enrichment >30), located within 2 kb of regions enriched in hESC H3K4me1 and H3K27ac (ChIP seeding fold enrichment >10 for both modifications), and, to distinguish these elements from proximal promoters, we demanded that these regions do not overlap with hESC H3K4me3 (ChIP seeding fold enrichment >30); class II elements (2,287 regions): genomic regions with hESC p300 enrichment (ChIP seeding fold enrichment >30), located within 2 kb of regions enriched in hESC H3K27me3 (ChIP seeding fold enrichment >8). These regions were further required not to overlap with hESC H3K4me3 (ChIP seeding fold enrichment >30) or hESC H3K27ac (ChIP seeding fold enrichment >10). Class II→I elements (195 regions): class II elements (as determined in hESCs) which in hNECs acquired enrichment in H3K27ac (H3K27ac ChIP seeding fold enrichment >10, within 2 kb of p300 peaks defining class II elements).

In total, we identified 11,543 regions marked by p300 and H3K4me1 in hESCs, of which 1,639 did not contain H3K27ac, H3K27me3 or H3K4me3 enrichment. A total of 3,531 regions were enriched for p300, H3K4me1 and H3K4me3 (those generally corresponded to proximal promoters).

Please note that although our definition of class II elements does not use an H3K4me1 enrichment filter, about 55% of class II regions are enriched for H3K4me1 at ChIP seeding fold enrichment >10; when lower cutoff is allowed, the overlap is significantly more substantial. Thus, the vast majority, if not all, class II elements probably contains above-background levels of H3K4me1, as exemplified by the observation that class II elements with ChIP-seq H3K4me1 levels below the seeding fold enrichment >10 cutoff are still substantially enriched for H3K4me1 when assayed by ChIP-qPCR (see Supplementary Fig. 5, for example, *CHD2*, *EPHA4*, *GPR19*, *ADRA2A*, *KLF5*, *EML1* regions).

**Other sequencing data analyses.** Average ChIP-seq signal profiles around the centre of p300-enriched regions were generated with the Sitepro tool, part of the Cistrome Analysis pipeline (<http://cistrome.dfci.harvard.edu/ap/>), using the corresponding WIG files generated with QuEST. Similarly, ChIP-seq signal profiles

were generated around gene TSS. For genes associated with the different classes of distal elements, each element was linked to its closest gene, based on the distance to TSS, and considering a maximum distance of 100 kb.

Average PhastCons scores profiles around the centre of p300-enriched regions were generated with the Conservation/Aggregate Datapoints tool, part of the Cistrome Analysis pipeline (<http://cistrome.dfci.harvard.edu/ap/>).

Distance between enhancers and their closest Ensembl gene TSS was calculated using PinkThing software (<http://pinkthing.cmbi.ru.nl/>) and Ensembl 52 assembly. With this information, it was possible to calculate the overall genomic distribution, based on distance to TSS, for the different enhancer groups and to assign enhancers to their closest genes.

Functional annotation of enhancers was obtained with GREAT (<http://great.stanford.edu/public/html/input.php>), using the Basal plus extension association rules and the whole human genome as background.

For RNA-seq data analysis, each enhancer was assigned to its closest gene based on distance to TSS considering a maximum distance of 100 kb, resulting in various gene groups each corresponding to an enhancer class (for example, class I, class II, class II→I). Statistical significance (*P*-values) of the difference in expression levels between different gene groups was calculated using two-sample one-sided Wilcoxon-test (R software, <http://www.r-project.org>). Paired or non-paired tests were performed when the same or different genes were compared, respectively. Box plots representing RPKM distribution were generated with R (<http://www.r-project.org>).

MRE-seq (methylation-sensitive restriction enzyme) data for hESCs was obtained from the GEO data set public repository under accession number GSM450236.

**In vitro enhancer reporter assays in hESCs and hNECs.** Representative class I, class II→I and class II elements (Supplementary Table 1) were cloned into a lentiviral vector (Sin-minTK-eGFP) in front of a minimal TK promoter driving GFP expression. hESC colonies were transduced with the appropriate lentiviruses and GFP fluorescence levels were subsequently monitored in undifferentiated hESCs, as well as in the course of hNEC differentiation (at day 1, 5 and 7 after induction of differentiation).

**Zebrafish reporter assays.** The biological relevance of the identified human enhancers was evaluated using Tol2 transposon-mediated transgenesis in zebrafish<sup>24</sup>. Selected human enhancers were PCR amplified and cloned in the pT2HE vector (gift from D. M. Kingsley), upstream of the *hsp70* promoter and eGFP. Tol2 transposase was *in vitro* transcribed using mMessage mMachine Sp6 kit (Ambion), according to the manufacturer's instructions. It is worth mentioning that the *hsp70* promoter independently drives robust and stable expression in the lens after 28–38 h.p.f.<sup>39</sup>. This lens signal is also observed when additional sequences are placed upstream of the minimal *hsp70* promoter, acting as a positive control for correct transgenesis. Vector DNA, with corresponding enhancers, and transposase RNA were mixed and injected in one-cell-stage zebrafish embryos as previously described. eGFP expression patterns were typically monitored at three different developmental times: 6–8 h.p.f., 10–14 h.p.f. and 24–28 h.p.f. According to ref. 24, using the described reporter assay method, 10–20% of the injected embryos are expected to display consistent and representative expression patterns. Because 50

embryos were typically injected, expression patterns were considered as representative for a given enhancer if displayed by at least 5–10 embryos within each batch (the remaining embryos typically showed a nonspecific or lack of fluorescence pattern). For those enhancers with identifiable and consistent expression patterns, a second set of injections (biological replicate) were performed for 50 additional embryos and in all cases similar results were obtained compared to the first injections.

Initial monitoring and embryo imaging were performed with a Leica M205 FA fluorescent stereoscope. High-resolution images presented in Fig. 4 were obtained with a Leica DM4500 B upright compound microscope.

Although live embryos were typically monitored and imaged, in order to obtain flat whole-embryo images, selected embryos were fixed and the yolk removed. Briefly, 24–28 h.p.f. embryos were dechorionated and transferred to 4% paraformaldehyde solution in PBS. After overnight rocking at 4 °C, fixed embryos were washed and stored in methanol at 20 °C until ready to use.

Specificity of our reporter assays was validated by assaying an extensive set of negative controls (Supplementary Table 4): (1) five class I elements; (2) four non-conserved genomic regions in proximity of four of the tested class II elements; (3) four human adult-tissue-specific enhancers that should not drive expression during early developmental stages; (4) three randomly selected intergenic non-conserved regions; (5) empty vector.

In addition, four selected class II elements were followed up to 5 days post-fertilization, together with their corresponding flanking non-conserved regions and additional negative controls. GFP patterns were monitored after 6 h.p.f., 24 h.p.f., 3 d.p.f. and 5 d.p.f. In these cases and for the class II elements, embryos showing specific patterns at the corresponding stage (for example, 6 h.p.f. for LEFTY2 and 24 h.p.f. for SOX2, EN1 and NKX2-1) were selected and their GFP patterns subsequently monitored. For the negative controls, once lens signal appeared (that is, transgenic embryos), such embryos were subsequently followed.

31. Boyer, L. A. *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947–956 (2005).
32. Furlan-Magaril, M., Rincon-Arango, H. & Recillas-Targa, F. Sequential chromatin immunoprecipitation protocol: ChIP-reChIP. *Methods Mol. Biol.* **543**, 253–266 (2009).
33. Ho, L. *et al.* An embryonic stem cell chromatin remodeling complex, esBAF, is an essential component of the core pluripotency transcriptional network. *Proc. Natl Acad. Sci. USA* **106**, 5187–5191 (2009).
34. Ieda, M. *et al.* Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. *Cell* **142**, 375–386 (2010).
35. Peng, J. C. *et al.* Jarid2/Jumonji coordinates control of PRC2 enzymatic activity and target gene occupancy in pluripotent cells. *Cell* **139**, 1290–1302 (2009).
36. Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R. & Young, R. A. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130**, 77–88 (2007).
37. Rahl, P. B. *et al.* c-Myc regulates transcriptional pause release. *Cell* **141**, 432–445 (2010).
38. Hargreaves, D. C., Horng, T. & Medzhitov, R. Control of inducible gene expression by signal-dependent transcriptional elongation. *Cell* **138**, 129–145 (2009).
39. Blechinger, S. R. *et al.* The heat-inducible zebrafish *hsp70* gene is expressed during normal lens development under non-stress conditions. *Mech. Dev.* **112**, 213–215 (2002).